

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2006, 23:33:41 ; Search time 27 Seconds
(without alignments)

2192.437 Million cell updates/sec

Title: US-10-659-055-1_COPY_51_766
Perfect score: 3877
Sequence: NTYRLKLSLRMISDHBYLY.....AHQHITYTHMSHPIKQCFSLP 716

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0\$
Maximum Match 100\$

Listing first 45 summaries

Database : Issued Patents AA:
1: /cgm2_6/prodata/1/iaa/5_COMB.pep:
2: /cgm2_6/prodata/1/iaa/6_COMB.pep:
3: /cgm2_6/prodata/1/iaa/H_COMB.pep:
4: /cgm2_6/prodata/1/iaa/PELTUS_COMB.pep:
5: /cgm2_6/prodata/1/iaa/RE_COMB.pep:
6: /cgm2_6/prodata/1/iaa/backfile1.pep:
*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	3877	100.0	755	4	PCT-US93-07923-3	Sequence 3, Appli
2	3877	100.0	759	4	PCT-US93-07923-2	Sequence 2, Appli
3	3877	100.0	766	1	US-08-230-491A-3	Sequence 3, Appli
4	3877	100.0	766	1	US-08-619-280A-3	Sequence 3, Appli
5	3877	100.0	766	1	US-08-940-591-3	Sequence 3, Appli
6	3877	100.0	766	2	US-09-794-236-1	Sequence 1, Appli
7	3877	100.0	766	2	US-09-265-606-3	Sequence 3, Appli
8	3877	100.0	775	2	US-09-949-016-0450	Sequence 1050, A
9	3871	99.8	766	2	US-10-002-593-6	Sequence 6, Appli
10	3871	99.8	766	2	US-09-949-016-6146	Sequence 6146, Ap
11	3871	99.8	766	2	US-10-423-114-6	Sequence 6, Appli
12	3867	99.7	766	2	US-09-518-550-27	Sequence 27, Appli
13	2948	76.0	593	4	PCT-US93-07923-11	Sequence 11, Appli
14	2138	55.1	760	1	US-08-230-191A-2	Sequence 2, Appli
15	2138	55.1	760	1	US-08-619-280A-2	Sequence 2, Appli
16	2138	55.1	760	1	US-08-940-391-2	Sequence 2, Appli
17	2138	55.1	760	2	US-09-265-606-2	Sequence 2, Appli
18	2136.5	55.1	761	2	US-09-518-550-66	Sequence 26, Appli
19	1229.5	31.7	547	2	US-09-949-016-8330	Sequence 8330, Ap
20	1217	31.4	796	2	US-09-976-674-5	Sequence 5, Appli
21	1198	30.9	743	2	US-10-363-337-4	Sequence 4, Appli
22	1196	30.8	706	2	US-09-976-674-41	Sequence 41, Appli
23	1152.5	29.7	691	2	US-09-976-674-43	Sequence 43, Appli
24	987	25.5	771	2	US-09-079-592-2	Sequence 2, Appli
25	963.5	24.9	771	2	US-09-462-284-2	Sequence 2, Appli
26	908	23.4	818	2	US-09-462-845-3	Sequence 3, Appli
27	908	23.4	818	2	US-10-402-312-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
PCT-US93-07923-3
; Sequence 3, Application PC/TUS9307923
; GENERAL INFORMATION:
; APPLICANT: Morimoto, Chikao
; SCHLOSSMAN, Stuart F.
; APPLICANT: Tanaka, Tohaki
; TITLE OF INVENTION: HUMAN CD26 AND METHODS FOR USE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: Wordperfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07923
; FILING DATE: 19930819
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/934,162
; FILING DATE: 21-AUG-1992
; APPLICATION NUMBER: 07/932,211
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00530/055002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 755
; TYPE: amino acid
; STRANDBEDNESS:
; TOPOLOGY: linear
; PCT-US93-07923-3

Query Match 100.0\$; Score 3877; DB 4; Length 755;
Best Local Similarity 100.0\$; Pred. No. 0\$;
Matches 716 ; Conservative 0\$; Mismatches 0\$; Indels 0\$; Gaps 0\$;

CLASSIFICATION: ;
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: 07/934,162
 FILING DATE: 21-AUG-1992
 APPLICATION NUMBER: 07/832,211
 FILING DATE: 06-FEB-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Fraser, Janis K.
 REGISTRATION NUMBER: 34,819
 REFERENCE DOCKET NUMBER: 00530/055002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 542-5070
 TELEFAX: (617) 542-8906
 TELEX: 200154
 SEQUENCE CHARACTERISTICS:
 LENGTH: 759
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 PCT-US93-07923-2

Query	Subject	Start	End	Length	Score	DB	Match Type
Qy	1 NYTRKLYSLWISDHEYLXKOENNLLVNEAYGNSSVFLNSTDFEGHSINDYSISPD	60	60	1	100.0*	4	Length 759;
Db	40 NYTRKLYSLWISDHEYLXKOENNLLVNEAYGNSSVFLNSTDFEGHSINDYSISPD	99	99	1	100.0*	4	Length 759;
Qy	61 GQFILLEBNYTKQRHRSYTAASYDIOLNKLRLITEREPIPNNTQWTSVPCHKLAYWN	120	120	1	100.0*	4	Length 759;
Db	100 GQFILLEBNYTKQRHRSYTAASYDIOLNKLRLITEREPIPNNTQWTSVPCHKLAYWN	159	159	1	100.0*	4	Length 759;
Qy	121 DIIYKIEPNLPSYRITWTGKEDIINYGITDVYEEVFSAYSALWSPNGFLAYAQFN	180	180	1	100.0*	4	Length 759;
Db	160 DIIYKIEPNLPSYRITWTGKEDIINYGITDVYEEVFSAYSALWSPNGFLAYAQFN	219	219	1	100.0*	4	Length 759;
Qy	181 TEPVPLIEYSFSDESLOQPKTRVPKAGAVNPVTFVNTDSLSSVTNATSIQTAP	240	240	1	100.0*	4	Length 759;
Db	220 TEPVPLIEYSFSDESLOQPKTRVPKAGAVNPVTFVNTDSLSSVTNATSIQTAP	279	279	1	100.0*	4	Length 759;
Qy	241 ASMLIGDHYLCDTWATQERISLWLRRIQNYSMIDCYDESSGRWNCLVARQHLEMST	300	300	1	100.0*	4	Length 759;
Db	280 ASMLIGDHYLCDTWATQERISLWLRRIQNYSMIDCYDESSGRWNCLVARQHLEMST	339	339	1	100.0*	4	Length 759;
Qy	301 TGWVGRFRSPSPHFTLDGNSYK1SNEEGRHICFOIDKDCPFTKGTVWIGIEAL	360	360	1	100.0*	4	Length 759;
Db	340 TGWVGRFRSPSPHFTLDGNSYK1SNEEGRHICFOIDKDCPFTKGTVWIGIEAL	399	399	1	100.0*	4	Length 759;
Qy	361 TSDLYYYISNEYKGMGPGRNLKYIKIQLSDYTCKTCLSCELNPEROQYSSVSKEAKYQ	420	420	1	100.0*	4	Length 759;
Db	400 TSDLYYYISNEYKGMGPGRNLKYIKIQLSDYTCKTCLSCELNPEROQYSSVSKEAKYQ	459	459	1	100.0*	4	Length 759;
Qy	421 RCGSGPLPLYTHSSNTDKGIRVLBONSALDKMLQNVMPBKCLDFTLNBTKEFYQML	480	480	1	100.0*	4	Length 759;
Db	460 RCGSGPLPLYTHSSNTDKGIRVLBONSALDKMLQNVMPBKCLDFTLNBTKEFYQML	519	519	1	100.0*	4	Length 759;
Qy	481 PHFDKSKCFCYPLLLDYAGPSQKADTVFRINWATYLASTNNIVASFDGGSGVYQGDKI	540	540	1	100.0*	4	Length 759;
Db	520 PHFDKSKCFCYPLLLDYAGPSQKADTVFRINWATYLASTNNIVASFDGGSGVYQGDKI	579	579	1	100.0*	4	Length 759;
Qy	541 MHAINRRLGTPEVDQTEAARQFSKMGFVDKRIAIWGSYGGYTTSMVLSGSGVFKCG	600	600	1	100.0*	4	Length 759;
Db	580 MHAINRRLGTPEVDQTEAARQFSKMGFVDKRIAIWGSYGGYTTSMVLSGSGVFKCG	639	639	1	100.0*	4	Length 759;
Qy	601 IAVAPVSRWEYDSVSYTERYNGLPLPBDNUHYRNSTMMSAENFKQVEYLILHGADDN	660	660	1	100.0*	4	Length 759;
Db	640 IAVAPVSRWEYDSVSYTERYNGLPLPBDNUHYRNSTMMSAENFKQVEYLILHGADDN	699	699	1	100.0*	4	Length 759;
Qy	661 VHFQSAQIISKALVDVGDFAMWTTDEDHGIASTAHQHLYTAHSFIKQCFSLP	716	716	1	100.0*	4	Length 759;
Db	700 VHFQSAQIISKALVDVGDFAMWTTDEDHGIASTAHQHLYTAHSFIKQCFSLP	755	755	1	100.0*	4	Length 759;
RESULT 2							
Sequence 2, Application PC/TUS9307923							
GENERAL INFORMATION:							
APPLICANT: Morimoto, Chikao							
APPLICANT: Schlossman, Stuart F.							
APPLICANT: Tanaka, Toshiaki							
TITLE OF INVENTION: HUMAN CDD6 AND METHODS FOR USE							
NUMBER OF SEQUENCES: 16							
CORRESPONDENCE ADDRESS:							
ADDRESS: Fish & Richardson							
STREET: 225 Franklin Street							
CITY: Boston							
STATE: Massachusetts							
COUNTRY: U.S.A.							
ZIP: 02110-2804							
COMPUTER READABLE FORM:							
MEDIUM TYPE: 3 1/2" Diskette, 1.44 Mb							
COMPUTER: IBM PS/2 Model 502 or 55SX							
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)							
SOFTWARE: WordPerfect (Version 5.0)							
CURRENT APPLICATION DATA:							
APPLICATION NUMBER: PCT/US93/07923							
FILING DATE: 19930819							
PCT/US93-07923-2							
Sequence 2, Application PC/TUS9307923							
GENERAL INFORMATION:							
APPLICANT: Morimoto, Chikao							
APPLICANT: Schlossman, Stuart F.							
APPLICANT: Tanaka, Toshiaki							
TITLE OF INVENTION: HUMAN CDD6 AND METHODS FOR USE							
NUMBER OF SEQUENCES: 16							
CORRESPONDENCE ADDRESS:							
ADDRESS: Fish & Richardson							
STREET: 225 Franklin Street							
CITY: Boston							
STATE: Massachusetts							
COUNTRY: U.S.A.							
ZIP: 02110-2804							
COMPUTER READABLE FORM:							
MEDIUM TYPE: 3 1/2" Diskette, 1.44 Mb							
COMPUTER: IBM PS/2 Model 502 or 55SX							
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)							
SOFTWARE: WordPerfect (Version 5.0)							
CURRENT APPLICATION DATA:							
APPLICATION NUMBER: PCT/US93/07923							
FILING DATE: 19930819							
PCT/US93-07923-2							
Sequence 2, Application PC/TUS9307923							
GENERAL INFORMATION:							
APPLICANT: Morimoto, Chikao							
APPLICANT: Schlossman, Stuart F.							
APPLICANT: Tanaka, Toshiaki							
TITLE OF INVENTION: HUMAN CDD6 AND METHODS FOR USE							
NUMBER OF SEQUENCES: 16							
CORRESPONDENCE ADDRESS:							
ADDRESS: Fish & Richardson							
STREET: 225 Franklin Street							
CITY: Boston							
STATE: Massachusetts							
COUNTRY: U.S.A.							
ZIP: 02110-2804							
COMPUTER READABLE FORM:							
MEDIUM TYPE: 3 1/2" Diskette, 1.44 Mb							
COMPUTER: IBM PS/2 Model 502 or 55SX							
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)							
SOFTWARE: WordPerfect (Version 5.0)							
CURRENT APPLICATION DATA:							
APPLICATION NUMBER: PCT/US93/07923							
FILING DATE: 19930819							
PCT/US93-07923-2							
Sequence 2, Application PC/TUS9307923							
GENERAL INFORMATION:							
APPLICANT: Morimoto, Chikao							
APPLICANT: Schlossman, Stuart F.							
APPLICANT: Tanaka, Toshiaki							
TITLE OF INVENTION: HUMAN CDD6 AND METHODS FOR USE							
NUMBER OF SEQUENCES: 16							
CORRESPONDENCE ADDRESS:							
ADDRESS: Fish & Richardson							
STREET: 225 Franklin Street							
CITY: Boston							
STATE: Massachusetts							
COUNTRY: U.S.A.							
ZIP: 02110-2804							
COMPUTER READABLE FORM:							
MEDIUM TYPE: 3 1/2" Diskette, 1.44 Mb							
COMPUTER: IBM PS/2 Model 502 or 55SX							
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)							
SOFTWARE: WordPerfect (Version 5.0)							
CURRENT APPLICATION DATA:							
APPLICATION NUMBER: PCT/US93/07923							
FILING DATE: 19930819							
PCT/US93-07923-2							
Sequence 2, Application PC/TUS9307923							
GENERAL INFORMATION:							
APPLICANT: Morimoto, Chikao							
APPLICANT: Schlossman, Stuart F.							
APPLICANT: Tanaka, Toshiaki							
TITLE OF INVENTION: HUMAN CDD6 AND METHODS FOR USE							
NUMBER OF SEQUENCES: 16							
CORRESPONDENCE ADDRESS:							
ADDRESS: Fish & Richardson							
STREET: 225 Franklin Street							
CITY: Boston							
STATE: Massachusetts							
COUNTRY: U.S.A.							
ZIP: 02110-2804							
COMPUTER READABLE FORM:							
MEDIUM TYPE: 3 1/2" Diskette, 1.44 Mb							
COMPUTER: IBM PS/2 Model 502 or 55SX							
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)							
SOFTWARE: WordPerfect (Version 5.0)							
CURRENT APPLICATION DATA:							
APPLICATION NUMBER: PCT/US93/07923							
FILING DATE: 19930819							
PCT/US93-07923-2							
Sequence 2, Application PC/TUS9307923							
GENERAL INFORMATION:							
APPLICANT: Morimoto, Chikao							
APPLICANT: Schlossman, Stuart F.							
APPLICANT: Tanaka, Toshiaki							
TITLE OF INVENTION: HUMAN CDD6 AND METHODS FOR USE							
NUMBER OF SEQUENCES: 16							
CORRESPONDENCE ADDRESS:							
ADDRESS: Fish & Richardson							
STREET: 225 Franklin Street							
CITY: Boston							
STATE: Massachusetts							
COUNTRY: U.S.A.							
ZIP: 02110-2804							
COMPUTER READABLE FORM:							
MEDIUM TYPE: 3 1/2" Diskette, 1.44 Mb							
COMPUTER: IBM PS/2 Model 502 or 55SX							
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)							
SOFTWARE: WordPerfect (Version 5.0)							
CURRENT APPLICATION DATA:							
APPLICATION NUMBER: PCT/US93/07923							
FILING DATE: 19930819							
PCT/US93-07923-2							
Sequence 2, Application PC/TUS9307923							
GENERAL INFORMATION:							
APPLICANT: Morimoto, Chikao							
APPLICANT: Schlossman, Stuart F.							
APPLICANT: Tanaka, Toshiaki							
TITLE OF INVENTION: HUMAN CDD6 AND METHODS FOR USE							
NUMBER OF SEQUENCES: 16							
CORRESPONDENCE ADDRESS:							
ADDRESS: Fish & Richardson							
STREET: 225 Franklin Street							
CITY: Boston							
STATE: Massachusetts							
COUNTRY: U.S.A.							
ZIP: 02110-2804							
COMPUTER READABLE FORM:							
MEDIUM TYPE: 3 1/2" Diskette, 1.44 Mb							
COMPUTER: IBM PS/2 Model 502 or 55SX							
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)							
SOFTWARE: WordPerfect (Version 5.0)							
CURRENT APPLICATION DATA:							
APPLICATION NUMBER: PCT/US93/07923							
FILING DATE: 19930819							
PCT/US93-07923-2							
Sequence 2, Application PC/TUS9307923							
GENERAL INFORMATION:							
APPLICANT: Morimoto, Chikao							
APPLICANT: Schlossman, Stuart F.							
APPLICANT: Tanaka, Toshiaki							
TITLE OF INVENTION: HUMAN CDD6 AND METHODS FOR USE							
NUMBER OF SEQUENCES: 16							
CORRESPONDENCE ADDRESS:							
ADDRESS: Fish & Richardson							
STREET: 225 Franklin Street							
CITY: Boston							
STATE: Massachusetts							
COUNTRY: U.S.A.							
ZIP: 02110-2804							
COMPUTER READABLE FORM:							
MEDIUM TYPE: 3 1/2" Diskette, 1.44 Mb							
COMPUTER: IBM PS/2 Model 502 or 55SX							
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)							
SOFTWARE: WordPerfect (Version 5.0)							
CURRENT APPLICATION DATA:							
APPLICATION NUMBER: PCT/US93/07923							
FILING DATE: 19930819							
PCT/US93-07923-2							
Sequence 2, Application PC/TUS9307923							
GENERAL INFORMATION:							
APPLICANT: Morimoto, Chikao							
APPLICANT: Schlossman, Stuart F.							
APPLICANT: Tanaka, Toshiaki							
TITLE OF INVENTION: HUMAN CDD6 AND METHODS FOR USE							
NUMBER OF SEQUENCES: 16							
CORRESPONDENCE ADDRESS:							
ADDRESS: Fish & Richardson							
STREET: 225 Franklin Street							
CITY: Boston							
STATE: Massachusetts							
COUNTRY: U.S.A.							
ZIP: 02110-2804							
COMPUTER READABLE FORM:							
MEDIUM TYPE: 3 1/2" Diskette, 1.44 Mb							
COMPUTER: IBM PS/2 Model 502 or 55SX							
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)							
SOFTWARE: WordPerfect (Version 5.0)							
CURRENT APPLICATION DATA:							
APPLICATION NUMBER: PCT/US93/07923							
FILING DATE: 19930819							
PCT/US93-07923-2							
Sequence 2, Application PC/TUS9307923							
GENERAL INFORMATION:							
APPLICANT: Morimoto, Chikao							
APPLICANT: Schlossman, Stuart F.							
APPLICANT: Tanaka, Toshiaki							
TITLE OF INVENTION: HUMAN CDD6 AND METHODS FOR USE							
NUMBER OF SEQUENCES: 16							
CORRESPONDENCE ADDRESS:							
ADDRESS: Fish & Richardson							
STREET: 225 Franklin Street							
CITY: Boston							
STATE: Massachusetts							
COUNTRY: U.S.A.							
ZIP: 02110-2804							
COMPUTER READABLE FORM:							
MEDIUM TYPE: 3 1/2" Diskette, 1.44 Mb							
COMPUTER: IBM PS/2 Model 502 or 55SX							
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)							
SOFTWARE: WordPerfect (Version 5.0)							
CURRENT APPLICATION DATA:							
APPLICATION NUMBER: PCT/US93/07923							
FILING DATE: 19930819							
PCT/US93-07923-2							
Sequence 2, Application PC/TUS9307923							
GENERAL INFORMATION:							
APPLICANT: Morimoto, Chikao							
APPLICANT: Schlossman, Stuart F.							
APPLICANT: Tanaka, Toshiaki							
TITLE OF INVENTION: HUMAN CDD6 AND METHODS FOR USE							
NUMBER OF SEQUENCES: 16							
CORRESPONDENCE ADDRESS:							
ADDRESS: Fish & Richardson							
STREET: 225 Franklin Street							
CITY: Boston							
STATE: Massachusetts							
COUNTRY: U.S.A.							
ZIP: 02110-2804							
COMPUTER READABLE FORM:							
MEDIUM TYPE: 3 1/2" Diskette, 1.44 Mb							
COMPUTER: IBM PS/2 Model 502 or 55SX							
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)							
SOFTWARE: WordPerfect (Version 5.0)							
CURRENT APPLICATION DATA:							
APPLICATION NUMBER: PCT/US93/07923							
FILING DATE: 19930819							
PCT/US93-07923-2							
Sequence 2, Application PC/TUS9307923							
GENERAL INFORMATION:							
APPLICANT: Morimoto, Chikao							
APPLICANT: Schlossman, Stuart F.							
APPLICANT: Tanaka, Toshiaki							
TITLE OF INVENTION: HUMAN CDD6 AND METHODS FOR USE							
NUMBER OF SEQUENCES: 16							
CORRESPONDENCE ADDRESS:							
ADDRESS: Fish & Richardson							
STREET: 225 Franklin Street							
CITY: Boston							
STATE: Massachusetts							
COUNTRY: U.S.A.							
ZIP: 02110-2804							
COMPUTER READABLE FORM:							
MEDIUM TYPE: 3 1/2" Diskette, 1.44 Mb							
COMPUTER: IBM PS/2 Model 502 or 55SX							
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)							
SOFTWARE: WordPerfect (Version 5.0)							
CURRENT APPLICATION DATA:							
APPLICATION NUMBER: PCT/US93/07923							
FILING DATE: 19930819							
PCT/US93-07923-2							
Sequence 2, Application PC/TUS9307923							
GENERAL INFORMATION:							
APPLICANT: Morimoto, Chikao							
APPLICANT: Schlossman, Stuart F.							
APPLICANT: Tanaka, Toshiaki							
TITLE OF INVENTION: HUMAN CDD6 AND METHODS FOR USE							
NUMBER OF SEQUENCES: 16							
CORRESPONDENCE ADDRESS:							
ADDRESS: Fish & Richardson							
STREET: 225 Franklin Street							
CITY: Boston							
STATE: Massachusetts							
COUNTRY: U.S.A.							
ZIP: 02110-2804							
COMPUTER READABLE FORM:							
MEDIUM TYPE: 3 1/2" Diskette, 1.44 Mb							
COMPUTER: IBM PS/2 Model 502 or 55SX							
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)							
SOFTWARE: WordPerfect (Version 5.0)							
CURRENT APPLICATION DATA:							
APPLICATION NUMBER: PCT/US93/07923							
FILING DATE: 19930819							
PCT/US93-07923-2							
Sequence 2, Application PC/TUS9307923							
GENERAL INFORMATION:							
APPLICANT: Morimoto, Chikao							
APPLICANT: Schlossman, Stuart F.							
APPLICANT: Tanaka, Toshiaki							
TITLE OF INVENTION: HUMAN CDD6 AND METHODS FOR USE							
NUMBER OF SEQUENCES: 16							
CORRESPONDENCE ADDRESS:							
ADDRESS: Fish & Richardson							
STREET: 225 Franklin Street							
CITY: Boston							
STATE: Massachusetts							
COUNTRY: U.S.A.							
ZIP: 02110-2804							
COMPUTER READABLE FORM:							
MEDIUM TYPE: 3 1/2" Diskette, 1.44 Mb							
COMPUTER: IBM PS/2 Model 502 or 55SX							
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)							
SOFTWARE: WordPerfect (Version 5.0)							
CURRENT APPLICATION DATA:							
APPLICATION NUMBER: PCT/US93/07923							
FILING DATE: 19930819							
PCT/US93-07923-2							
Sequence 2, Application PC/TUS9307923							
GENERAL INFORMATION:							
APPLICANT: Morimoto, Chikao							
APPLICANT: Schlossman, Stuart F.							
APPLICANT: Tanaka, Toshiaki							
TITLE OF INVENTION: HUMAN CDD6 AND METHODS FOR USE							
NUMBER OF SEQUENCES: 16							
CORRESPONDENCE ADDRESS:							
ADDRESS: Fish & Richardson							
STREET: 225 Franklin Street							
CITY: Boston							
STATE: Massachusetts							
COUNTRY: U.S.A.							
ZIP: 02110-2804							
COMPUTER READABLE FORM:							
MEDIUM TYPE: 3 1/2" Diskette, 1.44 Mb							
COMPUTER: IBM PS/2 Model 502 or 55SX							
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)							
SOFTWARE: WordPerfect (Version 5.0)							
CURRENT APPLICATION DATA:							
APPLICATION NUMBER: PCT/US93/07923							
FILING DATE: 19930819							
PCT/US93-07923-2							
Sequence 2, Application PC/TUS9307923							
GENERAL INFORMATION:							
APPLICANT: Morimoto, Chikao							
APPLICANT: Schlossman, Stuart F.							
APPLICANT: Tanaka, Toshiaki							
TITLE OF INVENTION: HUMAN CDD6 AND METHODS FOR USE							
NUMBER OF SEQUENCES: 16							
CORRESPONDENCE ADDRESS:							
ADDRESS: Fish & Richardson							
STREET: 225 Franklin Street							
CITY: Boston							
STATE: Massachusetts							
COUNTRY: U.S.A.							
ZIP: 02110-2804							
COMPUTER READABLE FORM:							
MEDIUM TYPE: 3 1/2" Diskette, 1.44 Mb							
COMPUTER: IBM PS/2 Model 502 or 55SX							
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)							
SOFTWARE: WordPerfect (Version 5.0)							
CURRENT APPLICATION DATA:							
APPLICATION NUMBER: PCT/US93/07923							
FILING DATE: 19930819							
PCT/US93-07923-2							
Sequence 2, Application PC/TUS9307923							
GENERAL INFORMATION:							
APPLICANT: Morimoto, Chikao							
APPLICANT: Schlossman, Stuart F.							
APPLICANT: Tanaka, Toshiaki							
TITLE OF INVENTION: HUMAN CDD6 AND METHODS FOR USE							
NUMBER OF SEQUENCES: 16							
CORRESPONDENCE ADDRESS:							
ADDRESS: Fish & Richardson							
STREET: 225 Franklin Street							
CITY: Boston							
STATE: Massachusetts							
COUNTRY: U.S.A.							
ZIP: 02110-2804							
COMPUTER READABLE FORM:							
MEDIUM TYPE: 3 1/2" Diskette, 1.44 Mb							
COMPUTER: IBM PS/2 Model 502 or 55SX							
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)							
SOFTWARE: WordPerfect (Version 5.0)							
CURRENT APPLICATION DATA:							
APPLICATION NUMBER: PCT/US93/07923							
FILING DATE: 19930819							
PCT/US93-07923-2							
Sequence 2, Application PC/TUS9307923							
GENERAL INFORMATION:							
APPLICANT: Morimoto, Chikao							
APPLICANT: Schlossman, Stuart F.							
APPLICANT: Tanaka, Toshiaki							
TITLE OF INVENTION: HUMAN CDD6 AND METHODS FOR USE							
NUMBER OF SEQUENCES: 16							
CORRESPONDENCE ADDRESS:							
ADDRESS: Fish & Richardson							
STREET: 225 Franklin Street							
CITY: Boston							
STATE: Massachusetts							
COUNTRY: U.S.A.							
ZIP: 02110-2804							
COMPUTER READABLE FORM:							
MEDIUM TYPE: 3 1/2" Diskette, 1.44 Mb							
COMPUTER: IBM PS/2 Model 502 or 55SX							
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)							
SOFTWARE: WordPerfect (Version 5.0)							
CURRENT APPLICATION DATA:							
APPLICATION NUMBER: PCT/US93/07923							
FILING DATE: 19930819							
PCT/US93-07923-2							
Sequence 2, Application PC/TUS9307923							
GENERAL INFORMATION:							
APPLICANT: Morimoto, Chikao							

RESULT 3

US-08-230-491A-3
Sequence 3, Application US/08230491A
Patent No. 5587299

GENERAL INFORMATION:

APPLICANT: Reetig, Wolfgang J.; Scanlan, Matthew J.;
Garin-Chesa, Pilar; Old, Lloyd J.

TITLE OF INVENTION: ISOLATED NUCLEAR ACID MOLECULE CODING FOR FIBROBLAST ACTIVATION PROTEIN AND USES

TITLE OF INVENTION: THEREOF

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:
ADDRESSEES: PELFE & LYNCH

STREET: 805 THIRD AVENUE

STATE: NEW YORK

COUNTRY: USA

ZIP: 10022

COMPUTER READABLE FORM:
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: WORDPERFECT - ASC II

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/230,491A
FILING DATE: 20-APRIL-1994
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5587299man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 330

TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9000
TELEFAX: (212) 638-3884

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 766 amino acids
TYPE: amino acid
TOPOLOGY: linear

US-08-230-491A-3

Query Match Score 3877; DB 1; Length 766;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0;
Indels 0; Gaps 0;

Qy 1 NTYRLKLYSRWISDEHYLQKENNIVFNAEYGNSSVFLENSDFEGHSINDYSISPD 60
Db 111 GQFILENTYKQWHRSYTASYDYLQNLQLITERIPNNTQWYTWSPNGTFLAYAQFND 120
51 NTYRLKLYSRWISDEHYLQKENNIVFNAEYGNSSVFLENSDFEGHSINDYSISPD 110

Qy 61 GQFILENTYKQWHRSYTASYDYLQNLQLITERIPNNTQWYTWSPNGTFLAYAQFND 120
Db 111 GQFILENTYKQWHRSYTASYDYLQNLQLITERIPNNTQWYTWSPNGTFLAYAQFND 170

Qy 121 DIVYKIEPNLPSYRITWTGKEDITYINGITDWYEEEVFSAYSALMWSPNGTFLAYAQFND 180
Db 171 DIVYKIEPNLPSYRITWTGKEDITYINGITDWYEEEVFSAYSALMWSPNGTFLAYAQFND 230

Qy 181 TEVPLIEYSPYSDESQYPKTQVTPYPKAGAVNPPTVKFFVNNTDSLSSVTNATSIQTAP 240
Db 231 TEVPLIEYSPYSDESQYPKTQVTPYPKAGAVNPPTVKFFVNNTDSLSSVTNATSIQTAP 290

Qy 241 ASMLJGDHYLCDVTWATQETISLOWLRTRONYSMDICDYEDESSGRMNLVARCHIEST 300
Db 291 ASMLJGDHYLCDVTWATQETISLOWLRTRONYSMDICDYEDESSGRMNLVARCHIEST 350

Qy 301 TGWVGRPSEPHFTLDGNSFYKLISNEGYRYCYFQDJKCOTPIKTGTWEIGIZL 360
Db 351 TGWVGRPSEPHFTLDGNSFYKLISNEGYRYCYFQDJKCOTPIKTGTWEIGIZL 410

Qy 361 TSDYLYXISMBKGMPGRNLYKIQSLDTKVTCLSCENPERCOYSYSFSKEAKYQOL 420
Db 51 NTYRLKLYSLRWISDEHYLQKENNIVFNAEYGNSSVFLENSDFEGHSINDYSISPD 110

Db 411 TSDYLYXISNEYKGMPGRNLYKIQSLDTKVTCLSCENPERCOYSYSFSKEAKYQOL 470
Qy 421 RCGSPGPPLPYTLLHSSYNDKSLRVLNQDMQMSKCLDFILNETKAYQML 480
Db 471 RCGSPGPPLPYTLLHSSYNDKSLRVLNQDMQMSKCLDFILNETKAYQML 530

Qy 481 PPHEDKSKKYPPLDYYAGPCSQADTVFLNATYLASTENITYASPDORGSGYQGDKI 540
Db 531 PPHEDKSKKYPPLDYYAGPCSQADTVFLNATYLASTENITYASPDORGSGYQGDKI 590

Qy 541 MHAINRRRLGFEVEQDQEAFQSMGFVNKRRAIWGNWGGYTTSVLFQGSGEVFKCG 600
Db 591 MHAINRRRLGFEVEQDQEAFQSMGFVNKRRAIWGNWGGYTTSVLFQGSGEVFKCG 650

Qy 601 JAVA PVS RME YYDSVYTYRMLGLPTPDENULDHYRNSTVNSRAENPKQVEYLHGTADDN 660
Db 651 JAVA PVS RME YYDSVYTYRMLGLPTPDENULDHYRNSTVNSRAENPKQVEYLHGTADDN 710

Qy 661 VHFQSAQI SKALDVGVDFQAMNTDEDHGIASTTAHQHLYTMSPHKQCFSLP 716
Db 711 VHFQSAQI SKALDVGVDFQAMNTDEDHGIASTTAHQHLYTMSPHKQCFSLP 766

RESULT 4
US-08-619-280A-3
; Sequence 3, Application US/08619280A
; Patent No. 5767242

GENERAL INFORMATION:
APPLICANT: Zimmermann, Rainer; Park, John E.;
TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN
TITLE OF INVENTION: ALPHA, AND USES THEREOF
NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:
ADDRESSEE: Felice & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/619-280A
FILING DATE: 18-MARCH-1996
CLASSIFICATION: 435
PRV APPLICATION DATA:
APPLICATION NUMBER: 08/230,491
FILING DATE: 20-APRIL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5767242man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5330.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-1884
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 766 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-619-280A-3

Query Match Score 3877; DB 1; Length 766;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0;
Indels 0; Gaps 0;

Qy	61	GOFILLEYNNVKQRHSYTAZYDYLINKROLITEERIPNNTQWNTSPVGHKLAYWN	120
Db	111	GOFILLEYNNVKQRHSYTAZYDYLINKROLITEERIPNNTQWNTSPVGHKLAYWN	170
Qy	121	DIVYKIEPNLPSRYTRTGKEDILLYNGTIDWYEEVFSAVSLAWNSPNGFLAYAQFND	180
Db	171	DIVYKIEPNLPSRYTRTGKEDILLYNGTIDWYEEVFSAVSLAWNSPNGFLAYAQFND	230
Qy	181	TEVILIEYSFSYSDLSLOQPKTVRVPYKAGAVNPYKFFVNTDSLSVVNTSATQTAP	240
Db	231	TEVILIEYSFSYSDLSQPKTVRVPYKAGAVNPYKFFVNTDSLSVVNTSATQTAP	290
Qy	241	ASMLIGDHYLCDVTWATERISQWLRLRIQNSYMDICDYDESSGRWNCLVARQHLEMST	300
Db	291	ASMLIGDHYLCDVTWATERISQWLRLRIQNSYMDICDYDESSGRWNCLVARQHLEMST	350
Qy	301	TGWNGRFRPSEPHFTLDGSNSFPYKLIISNEEGYRHICFYQIDKDCTFTTKGTMEVIGEAL	360
Db	351	TGWNGRFRPSEPHFTLDGSNSFPYKLIISNEEGYRHICFYQIDKDCTFTTKGTMEVIGEAL	410
Qy	361	TSDLYYYTSNEYKGMPGGRNLJKIQLSDPTYKVTCSCELNPERCQYVSFSKEAKYQY	420
Db	411	TSDLYYYTSNEYKGMPGGRNLJKIQLSDPTYKVTCSCELNPERCQYVSFSKEAKYQY	470
Qy	421	RCSCPGPLPYLTISSVNDKGRLVLEDNSALDKMLQNQMPSPKLLDFLILNETKFWQMIL	480
Db	471	RCSCPGPLPYLTISSVNDKGRLVLEDNSALDKMLQNQMPSPKLLDFLILNETKFWQMIL	530
Qy	481	PPHFDKSKKYPLLIIDVYAGPCSKORADTYFRLNWATYLASTENIIVASFDGRGSGKOGDKI	540
Db	531	PPHFDKSKKYPLLIIDVYAGPCSKORADTYFRLNWATYLASTENIIVASFDGRGSGXG9DKI	590
Qy	541	MHAIRNRRTGEVBDQIAARQFSKMGFDVNKRIAINGWSYGGYTSMLGSGSYFKCG	600
Db	591	MHAIRNRRTGEVBDQIAARQFSKMGFDVNKRATWGSYGGYTSMLGSGSYFKCG	650
Qy	601	IAYPVSRWEYDDSVTERYMLGPTEPDNLHYRNSTVMRSAENFKOVEYLILHGADDN	660
Db	651	IAYPVSRWEYDDSVTERYMLGPTEPDNLHYRNSTVMRSAENFKOVEYLILHGADDN	710
Qy	661	VHFQOSAQTSKALVYDVGYDFOAMWYTDDBDHGIASSTAHQHITYTHMSHFIKOCFSL	716
Db	711	VHFQOSAQTSKALVYDVGYDFOAMWYTDDBDHGIASSTAHQHITYTHMSHFIKOCFSL	766

RESULT 5
US-06-940-391-3
Sequence 3, Application US/08940391
Patent No. 5965373
GENERAL INFORMATION:
APPLICANT: Zimmerman, Rainier; Park, John E.;
APPLICANT: Rettig, Wolfgang; Old, Lloyd J.
TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION
TITLE OF INVENTION: PROTEIN ALPHA, AND USES THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Feile & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/940,391
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA:
CLASSIFICATION: 530

```

; APPLICATION NUMBER: 08/619,280
; FILING DATE: 18-MARCH-1996
; APPLICATION NUMBER: 08/230,491
; FILING DATE: 20-APRIL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, No. 5965373man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5330.1
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 3.
; SEQUENCE CHARACTERISTICS:
; LENGTH: 766 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-940-391-3

Query Match Score 3877; DB 1; Length 766;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTYRLKLYSLRWSIDHEYLKOENNLLVFNAYGNSSVFLENSTFDGFHSINDYSISSPD 60
Db 51 NTYRLKLYSLRWSIDHEYLQKNNLLVFNAYGNSSPLENSTFDGFHSINDYSISSPD 60
Qy 61 GQFILLENYVQWRHSTSASYDIYDLNKROLITER1PNTNTQWTTNSPVGHKLAYTWN 120
Db 111 GQPFLLENYVQWRHSTSASYDIYDLNKROLITER1PNTNTQWTTNSPVGHKLAYTWN 120
Qy 121 DIVKIEBNPLSPYRITWTGKD1IYNGITDWWYEVEVSAYSAWMSPNGTFLAYAQEND 180
Db 171 DIVKIEBNPLSPYRITWTGKD1IYNGITDWWYEVEVSAYSAWMSPNGTFLAYAQEND 230
Qy 181 TEVPLIEFSYSDESLQPKTYRVPYKPKAGAVNPTKFFEVVNTDSLSYYVTNATSIQTAP 240
Db 231 TEVPLIEFSYSDESLQPKTYRVPYKPKAGAVNPTKFFEVVNTDSLSYYVTNATSIQTAP 290
Qy 241 ASMLIGDHYLCDTWATERISIQLWLRRIQNYSMIDCDYDESSGRNCLVARQHTEMST 300
Db 291 ASMLIGDHYLCDTWATERISIQLWLRRIQNYSMIDCDYDESSGRNCLVARQHTEMST 350
Qy 301 TGWVGRFRSPSEPHFTLDGSNSFYKIIISNEEGYRHICYQIDKKDCTF1TKGMTEVIGEIAL 360
Db 351 TGWVGRFRSPSEPHFTLDGSNSFYKIIISNEEGYRHICYQIDKKDCTF1TKGMTEVIGEIAL 410
Qy 361 TSDYLYISNEYKGMGGRNLYKIQLSDPTKTYC1CBLNPRCQYYSVSFSKEAKYYQL 420
Db 411 TSDYLYISNEYKGMGGRNLYKIQLSDPTKTYC1CBLNPRCQYYSVSFSKEAKYYQL 470
Qy 421 RCGSPGLPLYTHSSVNDKGRLVLEDSALDKMLQNYQMPSKLDF1LNRTEKFWQMIL 480
Db 471 RCGSPGLPLYTHSSVNDKGRLVLEDSALDKMLQNYQMPSKLDF1LNRTEKFWQMIL 530
Qy 481 PPHFDKSCKYPLDVAAGPCSCOKADTVFRLNWATLASTEN1IVASFDRGSGSYQGDKI 540
Db 531 PPHFDKSCKYPLDVAAGPCSCOKADTVFRLNWATLASTEN1IVASFDRGSGSYQGDKI 590
Qy 541 MHAINRRGTFEVDQEAAQFSKMGFDVNKR1A1NGWSYGGYTTSMLGSGSYVKCG 600
Db 591 MHAINRRGTFEVDQEAAQFSKMGFDVNKR1A1NGWSYGGYTTSMLGSGSYVKCG 650
Qy 601 IAVPSERWEYDDSVYTERYMLG1PTBDNLHDYRNSTMRAENFKQVEYLLJHGADDN 660
Db 651 IAVPSERWEYDDSVYTERYMLG1PTBDNLHDYRNSTMRAENFKQVEYLLJHGADDN 710
Qy 661 VHFQQSAQTSKALVYDGVDFQAMWTTDEDHGIASSTAHQHTYTHMSPFIKOCFLSP 716
Db 711 VHFQQSAQTSKALVYDGVDFQAMWTTDEDHGIASSTAHQHTYTHMSPFIKOCFLSP 766

RESULT 6
US-09-784-216-1

```


RESULT 8
US-09-949-016-10450
; Sequence 10450, Application US/0949016
; Patent No. 6912339

; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 10450
; LENGTH: 775
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10450

RESULT 9
US-10-002-593-6
; Sequence 6, Application US/10002593
; Patent No. 6581198

; GENERAL INFORMATION:
; APPLICANT: Vanderbilt University
; ADDRESS: 3221瘾 University Station
; CITY: Nashville
; STATE: TN
; ZIP: 37235-6198
; COUNTRY: US
; TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOTENSIN CONVERTING ENZYME INHIBITOR AND VASOPERTIDE INHIBITOR ASSOCIATED ANGIOEDEMA
; FILE REFERENCE: Atty Docket No. 6581198 1242/48/2
; CURRENT APPLICATION NUMBER: US/10/002,593
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/244,524
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 6
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-002-593-6

Query Match 100.0%; Score 3877; DB 2; Length 775;
Best Local Similarity 100.0%; Prod. No. 0;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Query Match 99.8%; Score 3871; DB 2; Length 766;
Best Local Similarity 99.9%; Prod. No. 0;
Matches 715; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NTYRLKLSLRWISDHELYKQENNLVFAENGNSSVPLENSTDEFGHSINDYSISPD 60
Db 60 NTYRLKLSLRWISDHELYKQENNLVFAENGNSSVPLENSTDEFGHSINDYSISPD 119
Qy 61 QFILLENYKQWHRHTASYDIYLNKRQLITEERIPNTNTWSPVGHKLAYNN 120
Db 120 QFILLENYKQWHRHTASYDIYLNKRQLITEERIPNTNTWSPVGHKLAYNN 179
Qy 121 DIVYKIEPNLPSRITWTGKEDIIYNGTDWYEEEVAYSALWWSPNGTFLAYAQND 180
Db 180 DIVYKIEPNLPSRITWTGKEDIIYNGTDWYEEEVAYSALWWSPNGTFLAYAQND 239
Qy 181 TEVPLIEFSYSDSLLQPKTVRVPKAGAVNPVKFFVNTDSLSSVTNATSIQTAP 240
Db 240 TEVPLIEFSYSDSLLQPKTVRVPKAGAVNPVKFFVNTDSLSSVTNATSIQTAP 299
Qy 241 ASMLIGDHYLCDVWATQERISLQLWRIRQNSYMDICDYDESSGRMNCLVQHIEST 300
Db 300 ASMLIGDHYLCDVWATQERISLQLWRIRQNSYMDICDYDESSGRMNCLVQHIEST 359
Qy 301 TGWGRFRPSEPHFTLDGSNSPYK1ISNEGYRHICYQIDKKDCTFITKGTVWIGIAL 360

Db 360 TGWGRFRPSEPHFTLDGSNSFYK1ISNEGYRHICYQIDKKDCTFITKGTVWIGIAL 419
Qy 361 TSDLYLTISNEYKGMGPGRNLXKQLSPTYTKVTCLSCELNPERCQYVSFSKEAKYQL 420
Db 420 TSDLYLTISNEYKGMGPGRNLXKQLSPTYTKVTCLSCELNPERCQYVSFSKEAKYQL 479
Qy 421 RCGSPGLPLYTLHSYVDNGKLRLVLEDSALDKMLQNYOMPSKIKLDFILNETKFWYOML 480
Db 480 RCGSPGLPLYTLHSYVDNGKLRLVLEDSALDKMLQNYOMPSKIKLDFILNETKFWYOML 539
Qy 481 PPHFDKSKYKPLLLDVAGPCSKADTVFRLNWATYLASTENIVASFDGRSGYQGDKI 540
Db 540 PPHFDKSKYKPLLLDVAGPCSKADTVFRLNWATYLASTENIVASFDGRSGYQGDKI 599
Qy 541 MHA1NRRLGTFEVDQEAAAROFSDMGFVDNKRTAIWGMGYYTSMVLSGSSGVFCGG 600
Db 600 MHA1NRRLGTFEVDQEAAAROFSDMGFVDNKRTAIWGMGYYTSMVLSGSSGVFCGG 659
Qy 601 IAVAPVSRWEYDSVTERYMGILTPEPDNLHDYRNSTMRAENFKQVEYLJLHGADDN 660
Db 660 IAVAPVSRWEYDSVTERYMGILTPEPDNLHDYRNSTMRAENFKQVEYLJLHGADDN 719

Qy 661 VHQQSAQISKALVDPGVDQAMWYTDEDIASSTAHORIYTMISHFTIKQCFSLP 716
Db 720 VHQQSAQISKALVDPGVDQAMWYTDEDIASSTAHORIYTMISHFTIKQCFSLP 766

Qy	301 TCGWVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHYCYPQIDKKDCPFITKGOTWEVIGEAL 360	Db	231 TEVPLIEYSFYSDBSLQYPKTVRVYPKAGAVNPTVKFFVVNTDSLSSVTNATSIQTAP 290
Db	351 TCGWVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHYCYPQIDKKDCPFITKGOTWEVIGEAL 410	Qy	241 ASMLIGDHYLCDVTFATQBRISLQWLRRIONYSYMDICDYDESSGRWNCLVARDHIEMST 300
Qy	361 TSDLYLTISNEYKGMPGRNLXKQLSDYTKTCLSCELPERCOYSVSFSKEAKYQL 420	Db	291 ASMLIGDHYLCDVTFATQBRISLQWLRRIONYSYMDICDYDESSGRWNCLVARDHIEMST 350
Db	411 TSDLYLTISNEYKGMPGRNLXKQLSDYTKTCLSCELPERCOYSVSFSKEAKYQL 470	Qy	301 TGWVGRFRSEPHFTLDGNSFYKIIISNEEGYRHYCYPQIDKKDCPFITKGOTWEVIGEAL 360
Qy	421 RCGSGPGLPLYLHSSYNDKGVRVLENSALDKMLQNMQPSKCLDFIILNETKFWYQML 480	Db	351 TGWVGRFRSEPHFTLDGNSFYKIIISNEEGYRHYCYPQIDKKDCPFITKGOTWEVIGEAL 410
Db	471 RCGSGPGLPLYLHSSYNDKGVRVLENSALDKMLQNMQPSKCLDFIILNETKFWYQML 530	Qy	361 TSDLYLTISNEYKGMPGRNLXKQLSDYTKTCLSCELPERCOYSVSFSKEAKYQL 420
Qy	481 PPHEDFKSKYKPYLLDYYAGPCSQADTYFRLNWATYLASTENITIVASFDGRGSYQGDKI 540	Db	411 TSDLYLTISNEYKGMPGRNLXKQLSDYTKTCLSCELPERCOYSVSFSKEAKYQL 470
Db	531 PPHFDKSKYKPYLLDYYAGPCSQADTYFRLNWATYLASTENITIVASFDGRGSYQGDKI 590	Qy	421 RCGSGPGLPLYLHSSYNDKGVRVLENSALDKMLQNMQPSKCLDFIILNETKFWYQML 480
Qy	541 MHAIRNLGTFPEVEDOEARQFSKMGFVDNKRTAIWGSYGGTTSMYLGSGSVFRCG 600	Db	471 RCGSGPGLPLYLHSSYNDKGVRVLENSALDKMLQNMQPSKCLDFIILNETKFWYQML 530
Db	591 MHAIRNLGTFPEVEDOEARQFSKMGFVDNKRTAIWGSYGGTTSMYLGSGSVFRCG 650	Qy	481 PPHFDKSKYKPYLLDYYAGPCSQADTYFRLNWATYLASTENITIVASFDGRGSYQGDKI 540
Qy	601 IAVAPVSRWEYDSYTERMGLPTPBDNLHDYRSTMRAENFKQVEYLHTGADDN 660	Db	531 PPHFDKSKYKPYLLDYYAGPCSQADTYFRLNWATYLASTENITIVASFDGRGSYQGDKI 590
Db	651 IAVAPVSRWEYDSYTERMGLPTPBDNLHDYRSTMRAENFKQVEYLHTGADDN 710	Qy	541 MHAIRNLGTFPEVEDOEARQFSKMGFVDNKRTAIWGSYGGTTSMYLGSGSVFRCG 600
Qy	661 VHFOQSAQISKAVALDVGVDFQAMMTDEDHGIASSTAHQIYTMSHF1KQCFSLP 715	Db	591 MHAIRNLGTFPEVEDOEARQFSKMGFVDNKRTAIWGSYGGTTSMYLGSGSVFRCG 650
Db	711 VHFOQSAQISKAVALDVGVDFQAMMTDEDHGIASSTAHQIYTMSHF1KQCFSLP 766	Qy	601 IAVAPVSRWEYDSYTERMGLPTPBDNLHDYRSTMRAENFKQVEYLHTGADDN 660
Qy	US-09-949-016-6146 / Sequence 6146, Application US/0949016	Db	651 VHFOQSAQISKAVALDVGVDFQAMMTDEDHGIASSTAHQIYTMSHF1KQCFSLP 715
Qy	Patent No. 6812339	Db	711 VHFOQSAQISKAVALDVGVDFQAMMTDEDHGIASSTAHQIYTMSHF1KQCFSLP 766
<hr/>			
RESULT 10			
US-09-949-016-6146			
/ Sequence 6146, Application US/0949016			
GENERAL INFORMATION:			
APPLICANT: VENTER, J. Craig et al.			
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF			
FILE REFERENCE: CL001307			
CURRENT APPLICATION NUMBER: US 09/949, 016			
CURRENT FILING DATE: 2000-04-14			
PRIOR APPLICATION NUMBER: 60/241,755			
PRIOR FILING DATE: 2000-10-20			
PRIOR APPLICATION NUMBER: 60/237,768			
PRIOR FILING DATE: 2000-10-03			
PRIOR APPLICATION NUMBER: 60/231,498			
PRIOR FILING DATE: 2000-09-08			
NUMBER OF SEQ ID NOS: 207012			
SOFTWARE: FastSEQ for Windows Version 4.0			
SEQ ID NO: 6146			
LENGTH: 766			
TYPE: PRT			
ORGANISM: Human			
US-09-949-016-6146			
Query Match 99.8%; Score 3871; DB 2; Length 766;			
Best Local Similarity 99.9%; Pred. No. 0; Nismatches 0; Gaps 0;			
Matches 715; Conservative 0; Nismatches 1; Indels 0; Gaps 0;			
Qy	1 NTYRKLYSURWISDEHYLYKQENNIVFNAEYGNSSVFLLENSTDEFHSINDYSISPD 60	Qy	1 NTYRKLYSLRWSDEHYLYKQENNIVFNAEYGNSSVFLLENSTDEFHSINDYSISPD 60
Db	51 NTYRKLYSURWISDEHYLYKQENNIVFNAEYGNSSVFLLENSTDEFHSINDYSISPD 110	Db	51 NTYRKLYSLRWSDEHYLYKQENNIVFNAEYGNSSVFLLENSTDEFHSINDYSISPD 110
Qy	61 QGFIILENYVQWHRHSTSAYDYLQNLKRQLITEERIPNTNTWSPVGHKLAYWNN 120	Qy	61 QGFIILENYVQWHRHSTSAYDYLQNLKRQLITEERIPNTNTWSPVGHKLAYWNN 120
Db	111 QGFIILENYVQWHRHSTSAYDYLQNLKRQLITEERIPNTNTWSPVGHKLAYWNN 170	Db	111 QGFIILENYVQWHRHSTSAYDYLQNLKRQLITEERIPNTNTWSPVGHKLAYWNN 170
Qy	121 DIYVKEPLNPSPYRITWTGKEDIITYGTTWYEEEVFSAYSALMSPNGTFLAYAQND 180	Qy	121 DIYVKEPLNPSPYRITWTGKEDIITYGTTWYEEEVFSAYSALMSPNGTFLAYAQND 180
Db	171 DIYVKEPLNPSPYRITWTGKEDIITYGTTWYEEEVFSAYSALMSPNGTFLAYAQND 230	Db	171 DIYVKEPLNPSPYRITWTGKEDIITYGTTWYEEEVFSAYSALMSPNGTFLAYAQND 230
Qy	181 TEVPLIEYSFYSDSLEYQPKTVRVYPKAGAVNPTVKFFVVNTDSLSSVTNATSIQTAP 240	Qy	181 TEVPLIEYSFYSDSLEYQPKTVRVYPKAGAVNPTVKFFVVNTDSLSSVTNATSIQTAP 240

QY 181 TEVPLIEYSFYSDESLQYKPTKTRVVPYPKAGAVNPTVKFFVNTDSLSSVTNATSIQTAP 240
 Db 231 TEVPLIEYSFYSDESLQYKPTKTRVVPYPKAGAVNPTVKFFVNTDSLSSVTNATSIQTAP 290

QY 241 ASMLIGDHYLCDVTWATERISLOWLRIONSYMDICDYEDESSGRWNLVARQHIENST 300
 Db 291 ASMLIGDHYLCDVTWATERISLOWLRIONSYMDICDYEDESSGRWNLVARQHIENST 350

QY 301 TGMVGRFRPSEPHFTLDGNSFYK1ISNEEGYRHCYFQIDKDOCTFTKGTVWIGEAL 360
 Db 351 TGWVGRFRPSEPHFTLDGNSFYK1ISNEEGYRHCYFQIDKDOCTFTKGTVWIGEAL 410

QY 361 TSDLYIYISBEYKGMGPGRNLYK1QLSDYTKTVCUSCENPERQYVSFSREAKYQL 420
 Db 411 TSDLYIYISBEYKGMGPGRNLYK1QLSDYTKTVCUSCENPERQYVSFSREAKYQL 470

QY 421 RCGFGLPLTYLTLSVNDKLRVLEDNSAIDKMLQNYOMPSSKKLDFIINETKFWYQML 480
 Db 471 RCGFGLPLTYLTLSVNDKLRVLEDNSAIDKMLQNYOMPSSKKLDFIINETKFWYQML 530

QY 481 PPHFDKSCKYPLLIQDVAGPCSKOADTYFRLNWATYLASTENIIVASPDGRGSSYQDKI 540
 Db 531 PPHFDKSCKYPLLIQDVAGPCSKOADTYFRLNWATYLASTENIIVASPDGRGSSYQDKI 590

QY 541 MHA1NRRLGTFPEVEDQEARQFSKMGFDVNKR1AIWGSYGGVTSVTLGSSGGVFRCG 600
 Db 591 MHA1NRRLGTFPEVEDQEARQFSKMGFDVNKR1AIWGSYGGVTSVTLGSSGGVFRCG 650

QY 601 IAVAPSRWEYYDSVTERTYMLPTPEDNLHDYNSTMRAENFKQVEYLILHGADDN 660
 Db 651 IAVAPSRWEYYDSVTERTYMLPTPEDNLHDYNSTMRAENFKQVEYLILHGADDN 710

QY 661 VHFQOSAQISKALVDVGVDFOAMMYTDIEDHGIASSTAHQHYTMHSFTIKQFSLP 716
 Db 711 VHFQOSAQISKALVDVGVDFOAMMYTDIEDHGIASSTAHQHYTMHSFTIKQFSLP 766

RESULT 12
 US-09-518-550-27 Sequence No. 27, Application US/09518550

GENERAL INFORMATION:
 APPLICANT: TRAVIS, James
 APPLICANT: BANBULA, Agnieszka
 TITLE OF INVENTION: PROLYL PEPTIDASES AND METHODS OF USE
 FILE REFERENCE: 235-001901-01
 CURRENT APPLICATION NUMBER: US/09/518,550
 CURRENT FILING DATE: 2000-03-03
 PRIOR FILING DATE: 1999-03-05
 PRIOR APPLICATION NUMBER: PCT/US00/05551
 PRIOR FILING DATE: 2000-01-03
 NUMBER OF SEQ ID NOS: 48
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO: 27
 LENGTH: 766
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-518-550-27

Query Match 1 NTYRLKLYSLRWSIDPHEYLYKQENNVLVNAEYGNSSFLLENSTDFEGRHSINDYSISPD 60
 Matches 714; Conservative 99.7%; Score 3867; DB 2; Length 766;
 Best Local Similarity 99.7%; Pred. No. 0; N mismatches 1; Indels 0; Gaps 0;

QY 61 QQFILEXNTYKQWRHSTSAYDYLNLQRQLETEER1INNTQVTVSPVGHCLAYVNN 120
 Db 111 QQFILEXNTYKQWRHSTSAYDYLNLQRQLETEER1INNTQVTVSPVGHCLAYVNN 170

RESULT 13
 PCT-US93-07923-11
 ; Sequence 11, Application PC/TUS9307923
 ; GENERAL INFORMATION:
 ; APPLICANT: Morimoto, Chikao
 ; SCHLOSSMAN, Stuart F.
 ; APPLICANT: Tanaka, Toshiaki
 ; TITLE OF INVENTION: HUMAN CD26 AND METHODS FOR USE
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: U.S.A.
 ; ZIP: 02110-2804
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 ; COMPUTER: IBM PS/2 Model 502 or 55SX
 ; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
 ; SOFTWARE: WordPerfect (Version 5.0)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US93/07923
 ; FILING DATE: 19930819
 ; CLASSIFICATION:
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: 07/934,162
 ; FILING DATE: 21-AUG-1992
 ; APPLICATION NUMBER: 07/832,211
 ; FILING DATE: 06-FEB-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fraser, Janis K.

REGISTRATION NUMBER: 34,819
 REFERENCE/DOCKET NUMBER: 00530/055002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 542-5070
 TELEFAX: (617) 542-8906
 TELEX: 200154
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 593
 STRANDEDNESS: amino acid
 TOPOLOGY: linear
 FCT-US93-07923-11

Query Match 76.0%; Score 2948; DB 4; Length 593;
 Best Local Similarity 100.0%; Pred. No. 1.6e-252;
 Matches 543; Conservative 0; Mismatches 0;
 Indels 0; Gaps 0;

Qy 1 NTYRLKYSLRWISDHEYLQKENNLVFNABYGNSSVLENSTDFEGHSINDYSDSPD 60
 Db 51 NTYRLKYSLRWISDHEYLQKENNLVFNABYGNSSVLENSTDFEGHSINDYSDSPD 110

Qy 61 QOFILLENTYKQWRHSYTASYDIYDLNRQOLITERIPINTONTWSPYGHKLAYWN 120
 Db 111 QOFILLENTYKQWRHSYTASYDIYDLNRQOLITERIPINTONTWSPYGHKLAYWN 170

Qy 121 DIVVKLBPNLPSYRITWTGKEDITYNGIDTVYEEEFSSAYSALWSPNGETFLAYAQFND 180
 Db 171 DIVVKIEPNLPSYRITWTGKEDITYNGIDTVYEEEFSSAYSALWSPNGETFLAYAQFND 230

Qy 181 TEVPLIETSYFSDESQYKPTVTPPKAGAVNPYTKFFVNTDLSISVTNATSQTAP 240
 Db 231 TEVPLIETSYFSDESQYKPTVTPPKAGAVNPYTKFFVNTDLSISVTNATSQTAP 290

Db 241 ASMLIGCDHYLCDYTATQERISLQLRQIQLQSYNDICDYEDESSGRNCLVQRHIEMST 300
 Qy 291 ASMLIGCDHYLCDYTATQERISLQLRQIQLQSYNDICDYEDESSGRNCLVQRHIEMST 350

Qy 301 TGWGRFRPSPPHFTLDGSNSPYKILSNEEGYRHCYFQIDKKDCTFTIKTGTVWEIGEAL 360
 Db 351 TGWGRFRPSPPHFTLDGSNSPYKILSNEEGYRHCYFQIDKKDCTFTIKTGTVWEIGEAL 410

Qy 361 TSDYLIXYISNEYKGMPGRNLKYKQDSYTKVTCLSCELNPERCOYSSFSKEAKYQYL 420
 Db 411 TSDYLIXYISNEYKGMPGRNLKYKQDSYTKVTCLSCELNPERCOYSSFSKEAKYQYL 470

Qy 421 RCGSGPGLPLYTLHSSYNDKGKLRVLEDNSALDKMLQNVOMPSKKLDFIILNETKFWYQML 480
 Db 471 RCGSGPGLPLYTLHSSYNDKGKLRVLEDNSALDKMLQNVOMPSKKLDFIILNETKFWYQML 530

Qy 481 PHFDFDKSKCKTPLLDYAGPCSQKADTVFLRNWATYLASTENIIVASFDRGSGYQGDKI 540
 Db 531 PHFDFDKSKCKTPLLDYAGPCSQKADTVFLRNWATYLASTENIIVASFDRGSGYQGDKI 590

Qy 541 MHA 543
 Db 591 MHA 593

RESULT 14
 US-08-230-491A-2
 Sequence 2, Application US/08230491A
 Patent No. 5587299

GENERAL INFORMATION:
 APPLICANT: Retting, Wolfgang J.; Scanlan, Matthew J.;
 APPLICANT: Garin-Chesa, Pilar; Old, Lloyd J.
 TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE CODING FOR
 TITLE OF INVENTION: FIBROBLAST ACTIVATION PROTEIN, AND USES
 TITLE OF INVENTION: THEREOF
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: FELFE & LYNCH
 STREET: 805 THIRD AVENUE

CITY: NEW YORK STATE: NEW YORK
 COUNTRY: USA ZIP: 10022
 COMPUTER READABLE FORM:
 MEDIUM TYPE: DISKETTE 3.5 inch 1.2 MB STORAGE
 COMPUTER: IBM PS/2
 OPERATING SYSTEM: PC-DOS
 SOFTWARE: WORDPERFECT - ASC II
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/230,491A
 FILING DATE: 20-APRIL-1994
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Hanson, No. 5587299man D.
 REGISTRATION NUMBER: 30,946
 REFERENCE/DOCKET NUMBER: LUD 330
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 688-9200
 TELEFAX: (212) 838-3884
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 760 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 US-08-230-491A-2

Query Match 55.1%; Score 2138; DB 1; Length 760;
 Best Local Similarity 53.3%; Pred. No. 1.7e-180;
 Matches 383; Conservative 134; Mismatches 14; Gaps 7;
 Matches 383; Conservative 134; Mismatches 14; Gaps 7;

Qy 2 TYRKLKYSURWISDHEYLQK-ENNLVNAEYGNSSVLENSTDFEGHSIN--DYSIS 58
 Db 51 TFSYKTFEPNWISGEBYLHQSAADDNIVLYNIETGOSYTILSNRTRM---KSVNANSYGLS 106
 Qy 59 PDGGFTILLENNYVTKWRHSYTASYDIYDLNKQOLITERIPNTNTQWTVTSPVGHKLAYWN 118
 Db 107 PDRQTVLSDYSKLWRSYTTIYDLSNGEFTVRGNELPRPQYLCNSPVGSKLAYVN 166

Qy 119 NNDIYVKEIPNLPSYRITWTGKEDITYNGIDTVYEEEFSSAYSALWSPNGETFLAYAQF 178
 Db 167 QNNYTLKOREGDPFQITPGRENKIFNGKIDPQWYEEEMLPTKALWWSPNKFQFLAYAEF 226

Qy 179 NDTEVPLIETSYFSDESQYKPTVTPPKAGAVNPYTKFVNTDLSISVTNATSQT 238
 Db 227 NDKQTPVIAVSYCDE--QYPRTNIPYKAGANPVPVIRIIDTYTPAYGPQ---EVP 281

Qy 239 APASMLIGDHYLCDYTATQERISLQLRQIQLQSYNDICDYEDESSGRNCLVQRHIEM 298
 Db 282 VPAMIASDYYFSWLTWTDVFLQWLKRQNNTSICDFRDWQTMDCPKTQEHBEE 341

Qy 299 SRTGNGVGRFRPSEPHFTLDGSNSPYKILSNEEGYRHCYFQIDKKDCTFTIKTGTVEVIGIE 358
 Db 342 SRTGAGGFVSRVFSYDAISYKIFPSDDKGYHIIYKDTVNAIQTSKGWEAINIF 401

Qy 359 ALTDSDYLYISNEYKGMPGRNLKYKQDSYTKVTCLSCELNPERCOYSSFSKEAKYQYL 417
 Db 402 RVTQDSLFLYSSNNEPEYPPGRNTRISGYSPPSKKCYTCHLRERCQYTTASFSDAYC 461

Qy 418 YQRLCSGPGLPLYTLHSSYNDKGKLRVLEDNSALDKMLQNVOMPSKKLDFIILNETKFWYQ 477
 Db 462 YALVCYGPGLPISLHDGRTDQEIKLBEENKELNALKNIQLPKBEIKLDEVBLWYK 521

Qy 478 MILPHFEDSKKYYPLLDVYAGPCSQKADTVFLRNWATYLASTENIIVASFDRGSGYQ 537
 Db 522 MILPQFDSSKKCYVLLQYGGPCSQVSRSVFAVNNTSISLASKGMVIALDGRTAFQG 581

Qy 538 DKIMAHAINRLGTPEVEDQEAQFSKNGFVDNRRIATGWGSGYVTSMSVLGSQGV 597
 Db 582 DKLJAYVTKLGYVEVEDQITAFKPIEMGIDEKRIAWGMSGGYVSSLALASGTGLP 641

Qy 598 KCGIAVAPYRSRWEYDSYTERYMLGPTBDNUHYRNSTMRAENFKOVEYLIHGTA 657

Db 642 KCGIATAVPVSSWEYIASVYTERMGLPTKDDNLBHYKNSTVMARAEYFRNDVYLLIHGTA 701
 Qy 658 DDNVAHQSQSKALDVQDFAMWTFEDDGIASSTAHQHLYTAHSHPFKOCFSL 715
 Db 702 DDNHFQNSAQIAKLVNAQDFQAMWYSQDNHGL-SGUSTNHLYTHMFLQCPSL 758

RESULT 15
 US-08-619-280A-2
 Sequence 2, Application US/08619280A
 Patent No. 5,767,442
 GENERAL INFORMATION:
 APPLICANT: Zimmermann, Rainer; Park, John E.;
 ADDRESS: Bettig, Wolfgang; Old, Lloyd J.
 TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Feife & Lynch
 STREET: 805 Third Avenue
 CITY: New York City
 STATE: New York
 COUNTRY: USA
 ZIP: 10022
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
 COMPUTER: IBM PS/2
 OPERATING SYSTEM: PC-DOS
 SOFTWARE: Wordperfect
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/619,280A
 FILING DATE: 18-MARCH-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/230,491
 FILING DATE: 20-APRIL-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Hansen, No. 5767242 man D.
 REGISTRATION NUMBER: 30,946
 REFERENCE/DOCKET NUMBER: LUD 5330-1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 688-9200
 TELEFAX: (212) 638-3884
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 760 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 US-08-619-280A-2

Query Match 55.1%; Score 2118; DB 1; Length 760;
 Best Local Similarity 53.3%; Pred. No. 1.7e-180;
 Matches 383; Conservative 134; Mismatches 187; Indels 14; Gaps 7;

Qy 2 TYRLKLYSRWISDEHYLXQ-ENNIVLNAAEKGNSFLENSTFDEGHSHN--DYSIS 58
 Db 51 TFSKTKTFPWSQGEPYQKQADNIVNINETGQSYTNTTM--KSVNASVYCS 106

Qy 59 PDGQFTLLEYNYKQWRSHTASYDYLNLKROLITEERIPNNTQWWTSPVGHKLAVYW 118
 Db 107 PDRQFVYLBSDYSKLUWRYSTATYYDUSNGEVRGNELPRPIQYLCWSPVESKLAVY 166

Qy 119 NNDIYVKIEPNLPSYRITWKGEDIYNGSITDMMYEEFEVSAYSAIWWSPNGCFLAYQF 178
 Db 167 QNNIYVKQREGDPFPQTNGRENKIFINGLPDWYVEEMPLTPYALWWSPNGCFPLAYEF 226

Qy 179 NDEVPLIETFSYSDESQPKTRVPKAGAWNPYKFVNTDSVSSVNATSQT 238
 Db 227 NDKDIPVIAVSYCDB-QYPRTRNIPYKAGAKNPVYRIFLDTYAVVGQO -- EYP 281

Qy 239 APASMLGDYIQLCDVWTBATOBRISLQWLRQIYNSYMDICDYDESSGRWNCLVARQHLEM 298
 Db 282 VPAMIASDYYFPMNLJTWIDERVCLQWLKVQNVSVLSICDFREDWQTCPTQEHIEE 341

Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2006, 23:57:48 ; Search time 67 Seconds
(without alignments)

4465.161 Million cell updates/sec

Title: US-10-659-055-1_COPY_51_766

Perfect score: 3877

Sequence: 1_NTRKLXSLRNSDHEVLY.....AHOIHYTHMSHPIKQCFSLP 716

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 20000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:
1: /cgn2_6/pcodata/1/pubpaas/US07_PUBCOMB.pep;*
2: /cgn2_6/pcodata/1/pubpaas/US09_PUBCOMB.pep;*
3: /cgn2_6/pcodata/1/pubpaas/US10_PUBCOMB.pep;*
4: /cgn2_6/pcodata/1/pubpaas/US10A_PUBCOMB.pep;*
5: /cgn2_6/pcodata/1/pubpaas/US10B_PUBCOMB.pep;*
6: /cgn2_6/pcodata/1/pubpaas/US11_PUBCOMB.pep;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3877	100.0	766	3 US-09-265-606-3	Sequence 3, Appli
2	3877	100.0	766	3 US-09-293-959-1	Sequence 1, Appli
3	3877	100.0	766	4 US-10-295-027-590	Sequence 590, App
4	3877	100.0	766	5 US-10-952-459-18	Sequence 18, Appli
5	3877	100.0	766	5 US-10-631-457-565	Sequence 565, App
6	3871	99.8	766	4 US-10-002-533-6	Sequence 6, Appli
7	3871	99.8	766	4 US-10-165-603-7	Sequence 7, Appli
8	3871	99.8	766	4 US-10-123-714-6	Sequence 6, Appli
9	3871	99.8	766	4 US-10-295-027-922	Sequence 922, App
10	3871	99.8	766	4 US-10-794-899-41	Sequence 41, Appli
11	3871	99.8	766	6 US-11-041-674-6	Sequence 6, Appli
12	3867	99.7	766	5 US-10-476-664-147	Sequence 147, App
13	3867	99.7	766	5 US-10-770-712-1	Sequence 1, Appli
14	3358.5	86.6	767	5 US-10-770-712-2	Sequence 2, Appli
15	3347.5	86.3	767	4 US-10-165-603-4	Sequence 4, Appli
16	3347.5	86.3	767	4 US-10-794-899-38	Sequence 38, Appli
17	3340	86.1	760	5 US-10-631-457-190	Sequence 1390, Ap
18	3324	85.7	760	5 US-10-770-712-3	Sequence 3, Appli
19	2145	55.3	760	5 US-10-723-860-4171	Sequence 4171, Ap
20	2138	55.1	760	3 US-09-265-606-2	Sequence 2, Appli
21	2138	55.1	760	4 US-10-177-293-136	Sequence 136, App
22	2138	55.1	760	4 US-10-301-822-55	Sequence 55, Appli
23	2138	55.1	760	5 US-10-884-070-A3	Sequence 13, Appli
24	1276.5	32.9	504	4 US-10-072-012-863	Sequence 863, App
25	1221	31.5	228	5 US-10-476-264-114	Sequence 64, App
26	1217	31.4	746	5 US-10-476-364-69	Sequence 69, App
27	1217	31.4	746	5 US-10-476-264-70	Sequence 70, Appli

ALIGNMENTS

RESULT 1
US-09-265-606-3
; Sequence 3, Application US/09265606
; Patent No. US2002003489A1
; GENERAL INFORMATION:
; APPLICANT: Zimmermann, Rainer; Park, John E.;
; REITIG, Wolfgang Old, Lloyd J.
; TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN
; TITLE OF INVENTION: ALPHA, AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZEP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
; COMPUTER: IBM PS/2
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/265-606
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/619-280
; FILING DATE: 18-MARCH-1996
; APPLICATION NUMBER: 08/230,491
; FILING DATE: 20-APRIL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. US20020034789Alman D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5330.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-8884
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 766 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-09-265-606-3
Query Match Score 3877; DB 3; Length 766;
Best Local Similarity 100.0%; Pred. No. 5, 8e-312;
Matches 716; Conservations 0; Mismatches 0;
Indels 0; Gaps 0;

Qy 1 NTYRLKLYSLRWSIDHEYLYKQENNLILVFNAYEYNSSVFLNSTDEFCHSINDYSISPD 0

Db	51 NTYRLKLYSLRWSIDHELYXQENNILVNAEYGNSSVFLENSTDEFGHSINDYSISPD	110	Db	51 NTYRLKLYSLRWSIDHELYXQENNILVNAEYGNSSVFLENSTDEFGHSINDYSISPD	110
Qy	61 QFILELYNTVKQWHSYTASYDIYDNLNRQLITEERIPNNTQWTVSPVGHKLAYWN 120		Qy	61 GOFILLENYVKQRHHSYTASYDIYDNLNRQLITEERIPNNTQWTVSPVGHKLAYWN 120	
Db	111 QFILELYNTVKQWHSYTASYDIYDNLNRQLITEERIPNNTQWTVSPVGHKLAYWN 170		Db	111 GOFILLENYVKQRHHSYTASYDIYDNLNRQLITEERIPNNTQWTVSPVGHKLAYWN 170	
Qy	121 DIVYKIEPNLPSYRITWTGKEDITYNGTDMWYEEFVSAYSALMWSPNGTFLAYAQEND 180		Qy	121 DIVYKIEPNLPSYRITWTGKEDITYNGTDMWYEEFVSAYSALMWSPNGTFLAYAQEND 180	
Db	171 DIVYKIEPNLPSYRITWTGKEDITYNGTDMWYEEFVSAYSALMWSPNGTFLAYAQEND 230		Db	171 DIVYKIEPNLPSYRITWTGKEDITYNGTDMWYEEFVSAYSALMWSPNGTFLAYAQEND 230	
Qy	181 TEVPLIEYSFSYSDESLQPKTKVRPAGAVNPVFKFFVNTDSLSSVTNATS1QITAP 240		Qy	181 TEVPLIEYSFSYSDESLQPKTKVRPAGAVNPVFKFFVNTDSLSSVTNATS1QITAP 240	
Db	231 TEVPLIEYSFSYSDESLQPKTKVRPAGAVNPVFKFFVNTDSLSSVTNATS1QITAP 290		Db	231 TEVPLIEYSFSYSDESLQPKTKVRPAGAVNPVFKFFVNTDSLSSVTNATS1QITAP 290	
Qy	241 ASMLIGDHYLCDVTWATERISLOWLRTRIONSYMDICDYDESSGRMNLVARQHIENST 300		Qy	241 ASMLIGDHYLCDVTWATERISLOWLRTRIONSYMDICDYDESSGRMNLVARQHIENST 300	
Db	291 ASMLIGDHYLCDVTWATERISLOWLRTRIONSYMDICDYDESSGRMNLVARQHIENST 350		Db	291 ASMLIGDHYLCDVTWATERISLOWLRTRIONSYMDICDYDESSGRMNLVARQHIENST 350	
Qy	301 TGWYGRFRPSEPHFTLDGNSFSYKLTISNEEGYRHCYFQIDKKDCTFTIKGTMWVIGEAL 360		Qy	301 TGWYGRFRPSEPHFTLDGNSFSYKLTISNEEGYRHCYFQIDKKDCTFTIKGTMWVIGEAL 360	
Db	351 TGWYGRFRPSEPHFTLDGNSFSYKLTISNEEGYRHCYFQIDKKDCTFTIKGTMWVIGEAL 410		Db	351 TGWYGRFRPSEPHFTLDGNSFSYKLTISNEEGYRHCYFQIDKKDCTFTIKGTMWVIGEAL 410	
Qy	361 TSDVLYISNEYKGMPGGRNLYKQLSDYTKTCLSCELNPEROQYVSFSKEAKYQL 420		Qy	361 TSDVLYISNEYKGMPGGRNLYKQLSDYTKTCLSCELNPEROQYVSFSKEAKYQL 420	
Db	411 TSDVLYISNEYKGMPGGRNLYKQLSDYTKTCLSCELNPEROQYVSFSKEAKYQL 470		Db	411 TSDVLYISNEYKGMPGGRNLYKQLSDYTKTCLSCELNPEROQYVSFSKEAKYQL 470	
Qy	421 RCGFGLPLTYLHSSVNDKGLRVEDNSALDKMLQNVOMPSKLLDFIILNETKFWYQML 480		Qy	421 RCGFGLPLTYLHSSVNDKGLRVEDNSALDKMLQNVOMPSKLLDFIILNETKFWYQML 480	
Db	471 RCGFGLPLTYLHSSVNDKGLRVEDNSALDKMLQNVOMPSKLLDFIILNETKFWYQML 530		Db	471 RCGFGLPLTYLHSSVNDKGLRVEDNSALDKMLQNVOMPSKLLDFIILNETKFWYQML 530	
Qy	481 PPHFDKSKYKCYPLLLDYYAGPCSKADTYFRLNATYLASTENITIVASPDGRGSGYQDKI 540		Qy	481 PPHFDKSKYKCYPLLLDYYAGPCSKADTYFRLNATYLASTENITIVASPDGRGSGYQDKI 540	
Db	531 PPHFDKSKYKCYPLLLDYYAGPCSKADTYFRLNATYLASTENITIVASPDGRGSGYQDKI 590		Db	531 PPHFDKSKYKCYPLLLDYYAGPCSKADTYFRLNATYLASTENITIVASPDGRGSGYQDKI 590	
Qy	541 MHAINRRLGTFEVEDQIEARQFSKMGFDNKRTAIWGMSYGGVTSMVLGSSSGVFCRG 600		Qy	541 MHAINRRLGTFEVEDQIEARQFSKMGFDNKRTAIWGMSYGGVTSMVLGSSSGVFCRG 600	
Db	591 MHAINRRLGTFEVEDQIEARQFSKMGFDNKRTAIWGMSYGGVTSMVLGSSSGVFCRG 650		Db	591 MHAINRRLGTFEVEDQIEARQFSKMGFDNKRTAIWGMSYGGVTSMVLGSSSGVFCRG 650	
Qy	601 IAVAPVSRWEYDSUTTERYMGPLTPEDNLHDYNSTMRAENFKQVEYLILHGTDNN 660		Qy	601 IAVAPVSRWEYDSUTTERYMGPLTPEDNLHDYNSTMRAENFKQVEYLILHGTDNN 660	
Db	651 IAVAPVSRWEYDSUTTERYMGPLTPEDNLHDYNSTMRAENFKQVEYLILHGTDNN 710		Db	651 IAVAPVSRWEYDSUTTERYMGPLTPEDNLHDYNSTMRAENFKQVEYLILHGTDNN 710	
Qy	661 VHFQOSAQISKALVDGVDFQAMWYTDDEDHGIASSTAHOIYTHMSHTIKOCFSLP 716		Qy	661 VHFQOSAQISKALVDGVDFQAMWYTDDEDHGIASSTAHOIYTHMSHTIKOCFSLP 716	
Db	711 VHFQOSAQISKALVDGVDFQAMWYTDDEDHGIASSTAHOIYTHMSHTIKOCFSLP 766		Db	711 VHFQOSAQISKALVDGVDFQAMWYTDDEDHGIASSTAHOIYTHMSHTIKOCFSLP 766	

RESULT 2

US-09-933-959-1

; Sequence 1, Application US/09933959

; Publication No. US20030165489A1

; GENERAL INFORMATION:

; APPLICANT: Groumann, Eric

; APPLICANT: Lacroix, Jean-Silvain

; APPLICANT: Monod, Michel

; TITLE OF INVENTION: Method of Treating Rhinitis and Sinusitis

; FILE REFERENCE: 81985/276823

; CURRENT APPLICATION NUMBER: US/09/993,959

; CURRENT FILING DATE: 2001-11-27

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO: 1

; LENGTH: 766

; Query Match Score 3677; DB 3; Length 766;

; Best Local Similarity 100.0%; Pct. No. 5.8e-312;

; Matches 716; Conservatve 0; Mismatches 0; Indels 0; Gap 0;

; ORGANISM: Homo sapiens

US-09-933-959-1

RESULT 3

US-10-295-027-590

; Sequence 590, Application US/10295027

; Publication No. US20030232350A1

; GENERAL INFORMATION:

; APPLICANT: Afir, Daniel

; APPLICANT: Aziz, Natasha

; APPLICANT: Ginsberg, Wendy M.

; APPLICANT: Gish, Kurt C.

; APPLICANT: Glynn, Richard

; APPLICANT: Hevezsi, Peter A.

; APPLICANT: Mack, David H.

; APPLICANT: Murray, Richard

; APPLICANT: Watson, Susan R.

; APPLICANT: Bio-Biotechnology, Inc.

; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and

; TYPE: PCT; FILE REFERENCE: Methods of Screening for Modulators of Cancer

; CURRENT APPLICATION NUMBER: 018501-012500US

; CURRENT FILING DATE: 2002-11-13

; PRIORITY APPLICATION NUMBER: US/10/295,027

; PRIORITY FILING DATE: 2000-09-15

; PRIORITY APPLICATION NUMBER: US 60/350,666

; PRIORITY FILING DATE: 2001-11-13

; PRIORITY APPLICATION NUMBER: US 60/335,394

Qy

PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: US 60/332,464
PRIOR FILING DATE: 2001-11-21
PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: US 60/334,393
PRIOR FILING DATE: 2001-12-14
PRIOR APPLICATION NUMBER: US 60/347,211
PRIOR FILING DATE: 2002-01-08
PRIOR APPLICATION NUMBER: US 60/347,349
PRIOR FILING DATE: 2002-01-10
PRIOR APPLICATION NUMBER: US 60/355,250
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: US 60/356,714
PRIOR FILING DATE: 2002-02-13
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1386
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO: 590
TYPE: PRT
ORGANISM: Homo sapiens
US-10-295-027-590

Query Match 100.0%; Score 3877; DB 4; Length 766;
Best Local Similarity 100.0%; Pred. No. 5.8e-312;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTYRLKLSLRWISDHELYKQENNLVFNAYEGNNSVLENSTDEFGHSINDYSISPD 60
Db 51 NTYRLKLSLRWISDHELYKQENNLVFNAYEGNNSVLENSTDEFGHSINDYSISPD 110
Qy 61 QGFILEEYNTQKWRHSTSASYDYLNRQLITEERIPNTNTQWTVSPGHKLAYWNN 120
Db 111 QGFILEEYNTQKWRHSTSASYDYLNRQLITEERIPNTNTQWTVSPGHKLAYWNN 170
Qy 121 DYYKLEPNLPSRITWTGKEDITYITDWWYEEFVFSAYSALWSPNGTFLAYAQEND 180
Db 171 DYYKLEPNLPSRITWTGKEDITYITDWWYEEFVFSAYSALWSPNGTFLAYAQEND 230
Qy 181 TEVPLIEFSYSDSLOQPKTPKTVRPPKGAVNPVTKFFVNTDSLSSVTNATSIQTAP 240
Db 231 TEVPLIEFSYSDSLOQPKTPKTVRPPKGAVNPVTKFFVNTDSLSSVTNATSIQTAP 290
Qy 241 ASMLIGDHYLCDVTWATQRLSQLRIONYSYNDICDDESSRNCLVARHIEMST 300
Db 291 ASMLIGDHYLCDVTWATQRLSQLRIONYSYNDICDDESSRNCLVARHIEMST 350
Qy 301 TGWGRFRPSEPHFTLDGNSFYKITSNEEGYRHICYFQIDKKDCTFITKTGWEVIGEAL 360
Db 351 TGWGRFRPSEPHFTLDGNSFYKITSNEEGYRHICYFQIDKKDCTFITKTGWEVIGEAL 410
Qy 361 TSDLYYYISNEYKGMPGGRNLKYKQLSDTYKTCUCLSCELNPEROQYSYSFSKEAKYQL 420
Db 411 TSDLYYYISNEYKGMPGGRNLKYKQLSDTYKTCUCLSCELNPEROQYSYSFSKEAKYQL 470
Qy 421 RCGPGLPLTYLHSSVNDKGLRIVEDNSALDKMLQVNPMSKLLDFILNETKWTWQML 480
Db 471 RCGPGLPLTYLHSSVNDKGLRIVEDNSALDKMLQVNPMSKLLDFILNETKWTWQML 530
Qy 481 PPHEFDKSKYKPLLLDYAGPCSKQADTYFLNWTAYLASTENITIVASFDGRGSYQGDKI 540
Db 531 PPHEFDKSKYKPLLLDYAGPCSKQADTYFLNWTAYLASTENITIVASFDGRGSYQGDKI 590
Qy 541 MAHAINRRLGTFEVEDQIEAROFSKMGFDNKRIAIWGMNSGGYTSMVLGSGGVPKCG 600
Db 591 MAHAINRRLGTFEVEDQIEAROFSKMGFDNKRIAIWGMNSGGYTSMVLGSGGVPKCG 650
Qy 601 IAVAPVSRWEYYDSYTERMGLPPEVDNLDHYRNSTMSRAENFKQVEYLTLHTGADDN 660
Db 651 IAVAPVSRWEYYDSYTERMGLPPEVDNLDHYRNSTMSRAENFKQVEYLTLHTGADDN 710
Qy 661 VHFQOSAQISKALVGVDFQAMNYTDEDHGIASSTAQHITYTHMSHFIKQCFSLP 716

Db 711 VHFOOSAQISKALVGVDFQAMNYTDEDHGIASSTAQHITYTHMSHFIKQCFSLP 766

RESULT 4
US-10-952-459-18
Sequence 18, Application US/10952459
Publication No.: US20050074805A1
GENERAL INFORMATION:
APPLICANT: Kochan, Jarema Peter
APPLICANT: Martin, Mitchell Lee
APPLICANT: Rosinski, James Andrew
TITLE OF INVENTION: Specific Markers for Diabetes
FILE REFERENCE: 212/07US1
CURRENT APPLICATION NUMBER: US/10/952,459
CURRENT FILING DATE: 2004-09-28
NUMBER OF SEQ ID NOS: 34
SOFTWARE: Patentin version 3.2
SEQ ID NO: 18
LENGTH: 766
TYPE: PRT
ORGANISM: Homo sapiens
US-10-952-459-18

Query Match 100.0%; Score 3877; DB 5; Length 766;
Best Local Similarity 100.0%; Pred. No. 5.8e-312;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTYRLKLSLRWISDHELYKQENNLVFNAYEGNNSVLENSTDEFGHSINDYSISPD 60
Db 51 NTYRLKLSLRWISDHELYKQENNLVFNAYEGNNSVLENSTDEFGHSINDYSISPD 110
Qy 61 QGFILEEYNTQKWRHSTSASYDYLNRQLITEERIPNTNTQWTVSPGHKLAYWNN 120
Db 111 QGFILEEYNTQKWRHSTSASYDYLNRQLITEERIPNTNTQWTVSPGHKLAYWNN 170
Qy 121 DYYKLEPNLPSRITWTGKEDITYITDWWYEEFVFSAYSALWSPNGTFLAYAQEND 180
Db 171 DYYKLEPNLPSRITWTGKEDITYITDWWYEEFVFSAYSALWSPNGTFLAYAQEND 230
Qy 181 TEVPLIEFSYSDSLOQPKTPKTVRPPKGAVNPVTKFFVNTDSLSSVTNATSIQTAP 240
Db 231 TEVPLIEFSYSDSLOQPKTPKTVRPPKGAVNPVTKFFVNTDSLSSVTNATSIQTAP 290
Qy 241 ASMLIGDHYLCDVTWATQRLSQLRIONYSYNDICDDESSRNCLVARHIEMST 300
Db 291 ASMLIGDHYLCDVTWATQRLSQLRIONYSYNDICDDESSRNCLVARHIEMST 350
Qy 301 TGWGRFRPSEPHFTLDGNSFYKITSNEEGYRHICYFQIDKKDCTFITKTGWEVIGEAL 360
Db 351 TGWGRFRPSEPHFTLDGNSFYKITSNEEGYRHICYFQIDKKDCTFITKTGWEVIGEAL 410
Qy 361 TSDLYYYISNEYKGMPGGRNLKYKQLSDTYKTCUCLSCELNPEROQYSYSFSKEAKYQL 420
Db 411 TSDLYYYISNEYKGMPGGRNLKYKQLSDTYKTCUCLSCELNPEROQYSYSFSKEAKYQL 470
Qy 421 RCGPGLPLTYLHSSVNDKGLRIVEDNSALDKMLQVNPMSKLLDFILNETKWTWQML 480
Db 471 RCGPGLPLTYLHSSVNDKGLRIVEDNSALDKMLQVNPMSKLLDFILNETKWTWQML 530
Qy 481 PPHEFDKSKYKPLLLDYAGPCSKQADTYFLNWTAYLASTENITIVASFDGRGSYQGDKI 540
Db 531 PPHEFDKSKYKPLLLDYAGPCSKQADTYFLNWTAYLASTENITIVASFDGRGSYQGDKI 590
Qy 541 MAHAINRRLGTFEVEDQIEAROFSKMGFDNKRIAIWGMNSGGYTSMVLGSGGVPKCG 600
Db 591 MAHAINRRLGTFEVEDQIEAROFSKMGFDNKRIAIWGMNSGGYTSMVLGSGGVPKCG 650
Qy 601 IAVAPVSRWEYYDSYTERMGLPPEVDNLDHYRNSTMSRAENFKQVEYLTLHTGADDN 660
Db 651 IAVAPVSRWEYYDSYTERMGLPPEVDNLDHYRNSTMSRAENFKQVEYLTLHTGADDN 710
Qy 661 VHFOOSAQISKALVGVDFQAMNYTDEDHGIASSTAQHITYTHMSHFIKQCFSLP 716

Db	711 VHFQOSAQISKALVDVGVDFOAMNTYDDEHGIASTAHQIYTMMSHFIQCCSLP 765	651 IAVAPVSRWEYYDSVTERTYMLPPTPDNLHYRNSTMRAENFKQVEYLIIHGTAADDN 710
RESULT 5	Sequence 565, Application US/10631467	
;	Publication No. US20050208496A1	
;	GENERAL INFORMATION:	
;	APPLICANT: Genox Research Inc.	Sequence 6, Application US/10002593
;	TITLE OF INVENTION: Method for testing for bronchial asthma, or chronic obstructive p	Publication No. US20020137120A1
;	FILE REFERENCE: 3462-1005-00	GENERAL INFORMATION:
;	CURRENT APPLICATION NUMBER: US/10/631,467	APPLICANT: Vanderbilt University
;	CURRENT FILING DATE: 2003-07-31	APPLICANT: Brown, Nancy J.
;	PRIOR APPLICATION NUMBER: JP 2003-077212	TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOTENSIN CONVERTI
;	PRIOR FILING DATE: 2003-03-20	TITLE OF INVENTION: INHIBITOR AND VASOPEPTIDASE INHIBITOR ASSOCIATED ANGIOEDEMA
;	PRIOR APPLICATION NUMBER: JP 2002-229312	FILE REFERENCE: Atty Docket No. US20020137120A1 1242/48/2
;	PRIOR FILING DATE: 2002-08-06	CURRENT APPLICATION NUMBER: US/10/02,593
;	NUMBER OF SEQ ID NOS: 2086	CURRENT FILING DATE:
;	SOFTWARE: PatentIn version 3.1	PRIOR APPLICATION NUMBER: 60/244,524
;	SEQ ID NO: 565	PRIOR FILING DATE: 2000-10-31
;	LENGTH: 766	NUMBER OF SEQ ID NOS: 10
;	TYPE: PRT	SOFTWARE: PatentIn version 3.1
;	ORGANISM: Homo sapiens	SEQ ID NO: 6
;	US-10-631-467-565	LENGTH: 766
;	TYPE: PRT	TYPE: PRT
;	ORGANISM: Homo sapiens	ORGANISM: Homo sapiens
;	US-10-631-467-565	US-10-002-593-6
Query Match	100.0%; Score 3877; DB 5; Length 766;	
Best Local Similarity	100.0%; Pred. No. 5; 8e-312;	Query Match 99.8%; Score 3871; DB 4; Length 766;
Matches	716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	Best Local Similarity 99.9%; Prod. No. 1.e-311;
Qy	1 NTYRLKLYSLRWSIDSHELYKQENNLVNAEYGNSSYLENSTDFEFGHSINDYS1SPD 60	Matches 715; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db	51 NTYRLKLYSLRWSIDSHELYKQENNLVNAEYGNSSYLENSTDFEFGHSINDYS1SPD 110	Qy 1 NTYRLKLYSLRWSIDSHELYKQENNLVNAEYGNSSYLENSTDFEFGHSINDYS1SPD 60
Qy	61 GQFILLENYTKWKRHSYSTASYDIYDLNRQLITEERI PNNTOVTVWSVGHKLAYWN 120	Db 51 NTYRLKLYSLRWSIDSHELYKQENNLVNAEYGNSSYLENSTDFEFGHSINDYS1SPD 110
Db	111 GQFILLENYTKWKRHSYSTASYDIYDLNRQLITEERI PNNTOVTVWSVGHKLAYWN 170	Qy 61 GQFILLENYTKWKRHSYSTASYDIYDLNRQLITEERI PNNTOVTVWSVGHKLAYWN 120
Qy	121 DIVYKIEPNLPSYRITWTGKEDIYNGTIDWVYEEEVFSAYSALWWNSPNGTFLAYAQND 180	Db 111 GQFILLENYTKWKRHSYSTASYDIYDLNRQLITEERI PNNTOVTVWSVGHKLAYWN 170
Db	171 DIVYKIEPNLPSYRITWTGKEDIYNGTIDWVYEEEVFSAYSALWWNSPNGTFLAYAQND 230	Qy 121 DIVYKIEPNLPSYRITWTGKEDIYNGTIDWVYEEEVFSAYSALWWNSPNGTFLAYAQND 180
Qy	181 TEVPLIEYSFSDESLOQPKTVRVPYKAGAVNPTEKFVWNTDSLSSVTNATS1Q1TAP 240	Db 171 DIVYKIEPNLPSYRITWTGKEDIYNGTIDWVYEEEVFSAYSALWWNSPNGTFLAYAQND 230
Db	231 TEVPLIEYSFSDESLOQPKTVRVPYKAGAVNPTEKFVWNTDSLSSVTNATS1Q1TAP 290	Qy 181 TEVPLIEYSFSDESLOQPKTVRVPYKAGAVNPTEKFVWNTDSLSSVTNATS1Q1TAP 240
Qy	241 ASMLIGDHYLCDVWTQERISLQMLRTRIONYSYMDICDYDESSGRMNCVARQHIENST 300	Db 231 TEVPLIEYSFSDESLOQPKTVRVPYKAGAVNPTEKFVWNTDSLSSVTNATS1Q1TAP 290
Db	291 ASMLIGDHYLCDVWTQERISLQMLRTRIONYSYMDICDYDESSGRMNCVARQHIENST 350	Qy 241 ASMLIGDHYLCDVWTQERISLQMLRTRIONYSYMDICDYDESSGRMNCVARQHIENST 300
Qy	301 TGMVGRFRPSEPHFTLDGNSFYKLTISNEEGYRHICYFQDQDKCFTFKTGTVWIGIAL 360	Db 291 ASMLIGDHYLCDVWTQERISLQMLRTRIONYSYMDICDYDESSGRMNCVARQHIENST 350
Db	351 TGMVGRFRPSEPHFTLDGNSFYKLTISNEEGYRHICYFQDQDKCFTFKTGTVWIGIAL 410	Qy 301 TGMVGRFRPSEPHFTLDGNSFYKLTISNEEGYRHICYFQDQDKCFTFKTGTVWIGIAL 360
Qy	361 TSDI1YYISNEYKGMPGRNLKYKLQSLDTKVLSCBNPERQYVSFSKEAKYQL 420	Db 351 TGMVGRFRPSEPHFTLDGNSFYKLTISNEEGYRHICYFQDQDKCFTFKTGTVWIGIAL 410
Db	411 TSDI1YYISNEYKGMPGRNLKYKLQSLDTKVLSCBNPERQYVSFSKEAKYQL 470	Qy 361 TSDI1YYISNEYKGMPGRNLKYKLQSLDTKVLSCBNPERQYVSFSKEAKYQL 420
Qy	421 RCSGPGLPLYTLHSSVNDKLRVLEDNSALDKMLQNVOMPSKLLDFILNETKFWYQNL 480	Db 411 TSDI1YYISNEYKGMPGRNLKYKLQSLDTKVLSCBNPERQYVSFSKEAKYQL 470
Db	471 RCSGPGLPLYTLHSSVNDKLRVLEDNSALDKMLQNVOMPSKLLDFILNETKFWYQNL 530	Qy 421 RCSGPGLPLYTLHSSVNDKLRVLEDNSALDKMLQNVOMPSKLLDFILNETKFWYQNL 480
Qy	481 PPHEFDKSKKCYPLLDDVYAPCSQKADTVPLRNWATYLASTENITIVASPDGRGSGYQGKJ 540	Db 471 RCSGPGLPLYTLHSSVNDKLRVLEDNSALDKMLQNVOMPSKLLDFILNETKFWYQNL 530
Db	531 PPHEFDKSKKCYPLLDDVYAPCSQKADTVPLRNWATYLASTENITIVASPDGRGSGYQGKJ 590	Qy 481 PPHEFDKSKKCYPLLDDVYAPCSQKADTVPLRNWATYLASTENITIVASPDGRGSGYQGKJ 540
Qy	541 MHAINRRLGTFEVEDQIEARQFSKMGFDVNKRAIHWGSGYVTSMLGSSGGVPRCG 600	Db 531 PPHEFDKSKKCYPLLDDVYAPCSQKADTVPLRNWATYLASTENITIVASPDGRGSGYQGKJ 590
Db	591 MHAINRRLGTFEVEDQIEARQFSKMGFDVNKRAIHWGSGYVTSMLGSSGGVPRCG 650	Qy 541 MHAINRRLGTFEVEDQIEARQFSKMGFDVNKRAIHWGSGYVTSMLGSSGGVPRCG 600
Qy	601 IAVAPVSRWEYYDSVTERTYMLPPTPDNLHYRNSTMRAENFKQVEYLIIHGTAADDN 660	Db 591 MHAINRRLGTFEVEDQIEARQFSKMGFDVNKRAIHWGSGYVTSMLGSSGGVPRCG 650

Query 601 TAVAPVSRWEYDSSVTERYNGLPTPPDNLDHYRNSTMRAENFKQVEYLILGTADDN 660
 Db 651 TAVAPVSRWEYDSSVTERYNGLPTPPDNLDHYRNSTMRAENFKQVEYLILGTADDN 710

Query 661 VHFQQSAQISKALVDVDFQAMWYTDHDGIASTAHQIYTMSHPIKOCFLP 716
 Db 711 VHFQQSAQISKALVDVDFQAMWYTDHDGIASTAHQIYTMSHPIKOCFLP 766

RESULT 7
 US-10-165-603-7
 / Sequence 7, Application US/10165603
 / Publication No. US20030021792A1
 / GENERAL INFORMATION:
 / APPLICANT: Roben, Paul W.
 / TITLE OF INVENTION: TISSUE-SPECIFIC ENDOTHELIAL MEMBRANE
 / TITLE OF INVENTION: PROTEINS
 / FILE REFERENCE: TTECH 001A
 / CURRENT APPLICATION NUMBER: US/10/165,603
 / CURRENT FILING DATE: 2002-06-07
 / PRIOR APPLICATION NUMBER: 60/297,021
 / PRIOR FILING DATE: 2001-06-08
 / PRIOR APPLICATION NUMBER: 60/305,117
 / PRIOR FILING DATE: 2001-07-12
 / NUMBER OF SEQ ID NOS: 33
 / SOFTWARE: FastSEQ for Windows Version 4.0
 / SEQ ID NO 7
 / LENGTH: 766
 / TYPE: PRT
 / ORGANISM: Homo sapiens
 US-10-165-603-7

Query Match 99.8%; Score 3871; DB 4; Length 766;

Best Local Similarity 99.9%; Pred. No. 1.8e-311; Mismatches 0; Indels 0; Gaps 0;

Matches 715; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Query 1 NTYRKLYSLRWISDHEYLQKENNLIVPNAEYGNSSVLENSTDEFGHSINDYSISPD 60

Db 51 NTYRKLYSLRWISDHEYLQKENNLIVPNAEYGNSSVLENSTDEFGHSINDYSISPD 110

Query 61 GQFILELYNTYKQWRHSYTASYDIYDLNKROLITERIPNNTQWTVSPGHKLAYWNN 120

Db 111 GQFILELYNTYKQWRHSYTASYDIYDLNKROLITERIPNNTQWTVSPGHKLAYWNN 170

Query 121 DYYVKIEPNLPSYRITWTGEDIYINGITDWWYEEFVSEAYSALMWSPNGTFLAYAQFND 180

Db 171 DYYVKIEPNLPSYRITWTGEDIYINGITDWWYEEFVSEAYSALMWSPNGTFLAYAQFND 230

Query 181 TEVPLIEYSFSDESQYKPTVTPYKAGAVNPYKFFVNTDSLSVTNATSIQTAP 240

Db 231 TEVPLIEYSFSDESQYKPTVTPYKAGAVNPYKFFVNTDSLSVTNATSIQTAP 290

Query 241 ASMLIGDHYLCDVTAQERISLOWLRIONYNSMDICDYDESSGRWNLVAROHITEMST 300

Db 291 ASMLIGDHYLCDVTAQERISLOWLRIONYNSMDICDYDESSGRWNLVAROHITEMST 350

Query 301 TGWGRFRPSEPHFTLDGNSPKYKIIISNEEGYRHICYFOLDKDCDFITKTGWEVIGEAL 360

Db 351 TGWGRFRPSEPHFTLDGNSPKYKIIISNEEGYRHICYFOLDKDCDFITKTGWEVIGEAL 410

Query 361 TSDLYYISNEYKGMPGRNLKYQLSDYKTCUCLSCENPERCQYSYSFSKEARYXQL 420

Db 411 TSDLYYISNEYKGMPGRNLKYQLSDYKTCUCLSCENPERCQYSYSFSKEARYXQL 470

Query 421 RCGSPGLPLTHSSYNDKGLRVLEDNSALDKMLQVWQPSKCLDFIILNETKEFWYQML 480

Db 471 RCGSPGLPLTHSSYNDKGLRVLEDNSALDKMLQVWQPSKCLDFIILNETKEFWYQML 530

Query 481 PPHEFDKSKYKPLLLDQYAGPCSQKADTVFLRNWATYLASTENIIVASFDGRGSYQGDKI 540

Db 531 PPHEFDKSKYKPLLLDQYAGPCSQKADTVFLRNWATYLASTENIIVASFDGRGSYQGDKI 590

Query 541 MAIAINRRLGTFEVEDOIEAROFSKONGFVDNKRRAIWGNSYGGYVTSMVLJSGSGVFKCG 600
 Db 591 MAIAINRRLGTFEVEDOIEAROFSKONGFVDNKRRAIWGNSYGGYVTSMVLJSGSGVFKCG 650

Query 601 IAVAPVSRWEYDSSVTERYNGLPTPPDNLDHYRNSTMRAENFKQVEYLILGTADDN 660
 Db 651 IAVAPVSRWEYDSSVTERYNGLPTPPDNLDHYRNSTMRAENFKQVEYLILGTADDN 710

Query 661 VHFQQSAQISKALVDVDFQAMWYTDHDGIASTAHQIYTMSHPIKOCFLP 716
 Db 711 VHFQQSAQISKALVDVDFQAMWYTDHDGIASTAHQIYTMSHPIKOCFLP 766

RESULT 8
 US-10-423-714-6
 / Sequence 6, Application US/10423714
 / Publication No. US20030180828A1
 / GENERAL INFORMATION:
 / APPLICANT: Vanderbilt University
 / APPLICANT: Brown, Nancy J.
 / TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOTENSIN CONVERTING ENZYME INHIBITOR AND VASOPEPTIDASE INHIBITOR ASSOCIATED ANGIOEDEMA
 / FILE REFERENCE: Atty Docket No. US20030180828A1 1242/482
 / CURRENT APPLICATION NUMBER: US/10/423,714
 / CURRENT FILING DATE: 2003-04-25
 / PRIORITY APPLICATION NUMBER: 60/244,524
 / PRIORITY FILING DATE: 2000-10-31
 / NUMBER OF SEQ ID NOS: 10
 / SOFTWARE: Patentin version 3.1
 / SEQ ID NO 6
 / LENGTH: 766
 / TYPE: PRT
 / ORGANISM: Homo sapiens
 US-10-423-714-6

Query Match 99.8%; Score 3871; DB 4; Length 766;
 Best Local Similarity 99.9%; Pred. No. 1.8e-311; Mismatches 0; Indels 0; Gaps 0;
 Matches 715; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Query 1 NTYRKLYSLRWISDHEYLQKENNLIVPNAEYGNSSVLENSTDEFGHSINDYSISPD 60
 Db 51 NTYRKLYSLRWISDHEYLQKENNLIVPNAEYGNSSVLENSTDEFGHSINDYSISPD 110

Query 61 GQFILELYNTYKQWRHSYTASYDIYDLNRQLTTEERIPNNTQWTVSPGHKLAYWNN 120
 Db 111 GQFILELYNTVQQRHSYTASYDIYDLNRQLTTEERIPNNTQWTVSPGHKLAYWNN 170

Query 121 DYYVKIEPNLPSYRITWTGEDIYINGITDWWYEEFVSEAYSALMWSPNGTFLAYAQFND 180
 Db 171 DYYVKIEPNLPSYRITWTGEDIYINGITDWWYEEFVSEAYSALMWSPNGTFLAYAQFND 230

Query 181 TEVPLIEYSFSDESQYKPTVTPYKAGAVNPYKFFVNTDSLSVTNATSIQTAP 240
 Db 231 TEVPLIEYSFSDESQYKPTVTPYKAGAVNPYKFFVNTDSLSVTNATSIQTAP 290

Query 241 ASMLIGDHYLCDVTAQERISLOWLRIONYNSMDICDYDESSGRWNLVAROHITEMST 300
 Db 291 ASMLIGDHYLCDVTAQERISLOWLRIONYNSMDICDYDESSGRWNLVAROHITEMST 350

Query 301 TGWGRFRPSEPHFTLDGNSPKYKIIISNEEGYRHICYFOLDKDCDFITKTGWEVIGEAL 360
 Db 351 TGWGRFRPSEPHFTLDGNSPKYKIIISNEEGYRHICYFOLDKDCDFITKTGWEVIGEAL 410

Query 361 TSDLYYISNEYKGMPGRNLKYQLSDYKTCUCLSCENPERCQYSYSFSKEARYXQL 420
 Db 411 TSDLYYISNEYKGMPGRNLKYQLSDYKTCUCLSCENPERCQYSYSFSKEARYXQL 470

Query 421 RCGSPGLPLTHSSYNDKGLRVLEDNSALDKMLQVWQPSKCLDFIILNETKEFWYQML 480
 Db 471 RCGSPGLPLTHSSYNDKGLRVLEDNSALDKMLQVWQPSKCLDFIILNETKEFWYQML 530

Query 481 PPHEFDKSKYKPLLLDQYAGPCSQKADTVFLRNWATYLASTENIIVASFDGRGSYQGDKI 540
 Db 531 PPHEFDKSKYKPLLLDQYAGPCSQKADTVFLRNWATYLASTENIIVASFDGRGSYQGDKI 590

Db 531 PPHFDKSKYPLLLDVAGPCSQKADTVFRLNWTAYLASTENIVASSPGRGSGYQGDK1 590
 Qy 541 MHAINRRLGTEFEVEQIEARQFSKMGFVDNKRAIAGWSYGGVTSMLGSSGVFCG 600
 Db 591 MHAINRRLGTEFEVEQIEARQFSKMGFVDNKRAIAGWSYGG3VTSMVLGSSGVFCG 650
 Qy 601 IAVAVSRWEYDSYTERMGLPPDEDNDHYRNSTNSRAENFKQVYLLHTGADDN 660
 Db 651 IAVAVSRWEYDSYTERMGLPPDEDNDHYRNSTNSRAENFKQVYLLHTGADDN 710
 Qy 661 VHFQOSAISKALVQGVDFQAMWYTDDEHGIASSTAHQHITYTMSHFIKQCFSLP 716
 Db 711 VHFQOSAISKALVQGVDFQAMWYTDDEHGIASSTAHQHITYTMSHFIKQCFSLP 766

RESULT 9
 US-10-295-027-922
 ; Sequence 922, Application US/10295027
 ; Publication No. US20030232350A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Afar, Daniel
 ; APPLICANT: Aziz, Natasha
 ; APPLICANT: Ginsberg, Wendy M.
 ; APPLICANT: Gish, Kurt C.
 ; APPLICANT: Glynn, Richard
 ; APPLICANT: Hevezsi, Peter A.
 ; APPLICANT: Mack, David H.
 ; APPLICANT: Murray, Richard
 ; APPLICANT: Watson, Susan R.
 ; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
 ; FILE REFERENCE: 018501-0125001US
 ; CURRENT APPLICATION NUMBER: US/10/295, 027
 ; CURRENT FILING DATE: 2002-11-13
 ; PRIOR APPLICATION NUMBER: US/09/663,733
 ; PRIOR FILING DATE: 2000-09-15
 ; PRIOR APPLICATION NUMBER: US/09/666,666
 ; PRIOR FILING DATE: 2001-11-13
 ; PRIOR APPLICATION NUMBER: US/09/335,394
 ; PRIOR FILING DATE: 2001-11-15
 ; PRIOR APPLICATION NUMBER: US/09/332,464
 ; PRIOR FILING DATE: 2001-11-11
 ; PRIOR APPLICATION NUMBER: US/09/334,393
 ; PRIOR FILING DATE: 2001-11-29
 ; PRIOR APPLICATION NUMBER: US/09/340,376
 ; PRIOR FILING DATE: 2001-12-14
 ; PRIOR APPLICATION NUMBER: US/09/347,211
 ; PRIOR FILING DATE: 2001-01-08
 ; PRIOR FILING DATE: 2002-01-10
 ; PRIOR APPLICATION NUMBER: US/09/355,250
 ; PRIOR FILING DATE: 2002-05-08
 ; PRIOR APPLICATION NUMBER: US/09/356,714
 ; PRIOR FILING DATE: 2002-02-13
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 1386
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO: 922
 ; LENGTH: 766
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-295-027-922

Query Match 99.8%; Score 3871; DB 4; Length 766;
 Best Local Similarity 99.9%; Pred. No. 1..8e-311;
 Matches 715; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NYTRILKLYSLRWSIHDHELYKQENNIVNAEYGNSSVPLENSTDEFGHSINDYSISPD 60
 Db 51 NYTRILKLYSLRWSIHDHELYKQENNIVNAEYGNSSVPLENSTDEFGHSINDYSISPD 110
 Qy 61 GQFILENNYVQKWRHSTSASYDIYDLNRQLITEERIPNNTQWTVNSPVGHKLAYWNN 120
 Db 111 GQFILENNYVQKWRHSTSASYDIYDLNRQLITEERIPNNTQWTVNSPVGHKLAYWNN 170

Db 111 GQFILENNYVQKWRHSTSASYDIYDLNRQLITEERIPNNTQWTVNSPVGHKLAYWNN 170
 Qy 121 DRYVKIEBNLPSTRITWGKEDIIYNGTIDWYEEVSAYSAWWSSNGTFLAYAQFND 180
 Db 171 DRYVKIEBNLPSTRITWGKEDIIYNGTIDWYEEVSAYSAWWSSNGTFLAYAQFND 230
 Qy 181 TEVPLIEFSYSDSLOYPKTVRVPYPKAGAVNPVTKFFVNTDSSLSEVTNATSIQTAP 240
 Db 231 TEVPLIEFSYSDSLOYPKTVRVPYPKAGAVNPVTKFFVNTDSSLSEVTNATSIQTAP 290
 Qy 241 ASMLIGDHYLCDVWATERISLQWLRIQNYSYMIDCDYDESSGRNCLVARQHIENST 300
 Db 291 ASMLIGDHYLCDVWATERISLQWLRIQNYSYMIDCDYDESSGRNCLVARQHIENST 350
 Qy 301 TGWGRFRPSEPHFTLDGNSFYK1ISNEGYRHICYF01DKDCTF1TKGTWEVIGEAL 360
 Db 351 TGWGRFRPSEPHFTLDGNSFYK1ISNEGYRHICYF01DKDCTF1TKGTWEVIGEAL 410
 Qy 361 TSDLYTTSNEYKGMPGGRNLXK1QLSPYXTKVTCLSCELNPERCQYVSFSKAKYQOL 420
 Db 411 TSDLYTTSNEYKGMPGGRNLXK1QLDYTAKVTCLSCELNPERCQYVSFSKAKYQOL 470
 Qy 421 RCSCPGPLYLTHS SYNDKG1LVMJEDNSALDKMQLQNYQMPSKKCLDF1LNETKFWQML 480
 Db 471 RCSCPGPLYLTHS SYNDKG1LVMJEDNSALDKMQLQNYQMPSKKCLDF1LNETKFWQML 530
 Qy 481 PPHFDKSKYPLLLUDTAGPCSORADTVPLWATYLASTENIVASPDGRGSGYQGDKI 540
 Db 531 PPHFDKSKYPLLLUDTAGPCSORADTVPLWATYLASTENIVASPDGRGSGYQGDKI 590
 Qy 541 MHA1NRLRGTFEVDQ1EAAROFSKMGEVDNKR1ATMGWSYGGVTSMLGSGSVFKG 600
 Db 591 MHA1NRLRGTFEVDQ1EAAROFSKMGEVDNKR1ATMGWSYGGVTSMLGSGSVFKG 650
 Qy 601 IAVAPVSRWEYDSVTERYGM1PTPEDNLHYRNSTMRAENFKQVBYLJHGADDN 660
 Db 651 IAVAPVSRWEYDSVTERYGM1PTPEDNLHYRNSTMRAENFKQVBYLJHGADDN 710
 Qy 661 VHFQOSAQISKALVQGVDFQAMWYTDDEHGIASSTAHQHITYTMSHFIKQCFSLP 716
 Db 711 VHFQOSAQISKALVQGVDFQAMWYTDDEHGIASSTAHQHITYTMSHFIKQCFSLP 766

RESULT 10
 US-10-794-899-41
 ; Sequence 41, Application US/10794899
 ; Publication No. US20040146516A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Utah Ventures
 ; TITLE OF INVENTION: Lumen-Exposed Molecules and Methods Targeted Delivery
 ; FILE REFERENCE: 27110-715
 ; CURRENT APPLICATION NUMBER: US/10/794, 899
 ; CURRENT FILING DATE: 2004-03-05
 ; NUMBER OF SEQ ID NOS: 121
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO: 41
 ; LENGTH: 766
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-794-899-41

Query Match 99.8%; Score 3871; DB 4; Length 766;
 Best Local Similarity 99.9%; Pred. No. 1..8e-311;
 Matches 715; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NYTRILKLYSLRWSIHDHELYKQENNIVNAEYGNSSVPLENSTDEFGHSINDYSISPD 60
 Db 51 NYTRILKLYSLRWSIHDHELYKQENNIVNAEYGNSSVPLENSTDEFGHSINDYSISPD 110
 Qy 61 GQFILENNYVQKWRHSTSASYDIYDLNRQLITEERIPNNTQWTVNSPVGHKLAYWNN 120
 Db 111 GQFILENNYVQKWRHSTSASYDIYDLNRQLITEERIPNNTQWTVNSPVGHKLAYWNN 170

Qy 121 DIYVKLEPNLPSYRITWTGKEDITYINGITDVKFVYEEVFSAYSALWMSPNGTFLAYAQFND 180
 Db 171 DIYVKLEPNLPSYRITWTGKEDITYINGITDVKFVYEEVFSAYSALWMSPNGTFLAYAQFND 230
 Qy 181 TEVPLIEYSFYSDESLQYPKTRVTPPKAGAVNPTVKFVYNTDSLSVTNATSIOTIAP 240
 Db 231 TEVPLIEYSFYSDESLQYPKTRVTPPKAGAVNPTVKFVYNTDSLSVTNATSIOTIAP 290
 Qy 241 ASMLIGDHYLCDVWTATQERISLQMLRRIQNYSMDICDYESSGRNCLVARQHTEMST 300
 Db 291 ASMLIGDHYLCDVWTATQERISLQMLRRIQNYSMDICDYESSGRNCLVARQHTEMST 350
 Qy 301 TGDWGRFRPSPBHFITDGNSPYK1ISNEEGYRHICFQIDKDCDFITKGTVWIGIEAL 360
 Db 351 TGDWGRFRPSPBHFITDGNSPYK1ISNEEGYRHICFQIDKDCDFITKGTVWIGIEAL 410
 Qy 361 TSDLYYYISNEYKGMPGGRNLKYKIQLSDYTKVTCCLSCELLPERCOYVSFSKEAKYQL 420
 Db 411 RCGSGPSPFLYLTHSSYNDKGRLVERLEDNSALDKMLQNVQMPSKKDFILNETKFWQMIL 480
 Qy 421 RCGSGPSPFLYLTHSSYNDKGRLVERLEDNSALDKMLQNVQMPSKKDFILNETKFWQMIL 480
 Db 471 RCGSGPSPFLYLTHSSYNDKGRLVERLEDNSALDKMLQNVQMPSKKDFILNETKFWQMIL 530
 Qy 481 PPHFDKSCKPFLLDVYAGPCSQADTVFLNWATYLASTENIIVASPDGSGYQGDKI 540
 Db 531 PPHFDKSCKPFLLDVYAGPCSQADTVFLNWATYLASTENIIVASPDGSGYQGDKI 590
 Qy 541 MHAINRRLGTEPEVDQEAAQFSKMGFVDNKRIALWGMSYGGYTTSMNGSGSVEFKCG 600
 Db 591 MHAINRRLGTEPEVDQEAAQFSKMGFVDNKRIALWGMSYGGYTTSMNGSGSVEFKCG 650
 Qy 601 IAVAPYSRWEEYDSUTERMGLPTPEDNDHYRASTVMSRAENPKQVEYLILHGTDNN 660
 Db 651 IAVAPYSRWEEYDSUTERMGLPTPEDNDHYRASTVMSRAENPKQVEYLILHGTDNN 710
 Qy 661 VHFQOQAQISKALVGDFOAMMFTDEDGIGIASSTAHQHITYTMSPFIKOCFSLP 716
 Db 711 VHFQOQAQISKALVGDFOAMMFTDEDGIGIASSTAHQHITYTMSPFIKOCFSLP 766

RESULT 11
 US-11-041-674-6
 / Sequence 6, Application US/11041674
 / Publication No. US20050181168A1
 / GENERAL INFORMATION:
 / APPLICANT: Vanderbilt University
 / ATTORNEY: Brown, Nancy J.
 / TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOTENSIN CONVERTIN
 / FILE REFERENCE: Accy Docket No. 1242/48/2/2
 / CURRENT APPLICATION NUMBER: US/11/041,674
 / PRIORITY APPLICATION NUMBER: 2005-01-24
 / PRIORITY FILING DATE: 2000-10-31
 / PRIOR APPLICATION NUMBER: 10/002,593
 / PRIORITY FILING DATE: 2001-10-31
 / PRIOR APPLICATION NUMBER: 10/423,714
 / PRIORITY FILING DATE: 2003-4-25
 / NUMBER OF SEQ ID NOS: 10
 / SOFTWARE: PatentIn version 3.1
 / LENGTH: 766
 / ORGANISM: Homo sapiens
 US-11-041-674-6

Query Match 99.8%; Score 3871; DB 6; Length 766;
 Best Local Similarity 99.9%; Pred. No. 1..8e-311;
 Matches 715; Conservative 0; Mismatches -1; Indels 0; Gaps 0;

51 NTYRLKLSLWISDELYQENTNLVNAEYGNSSVLENSTDEFGHSDINDYSISPD 110
 61 CQFILEEYNTKQMRHSYTASYDIDLNRKOLITEERIPNTNTQWTTSIVGKLLAYWN 120
 111 CQFILEEYNTKQMRHSYTASYDIDLNRKOLITEERIPNTNTQWTTSIVGKLLAYWN 170
 121 DIYVKLEPNLPSYRITWTGEDIITNGITDVKFVYEEVFSAYSALWMSPNGTFLAYAQFND 180
 171 DIYVKLEPNLPSYRITWTGEDIITNGITDVKFVYEEVFSAYSALWMSPNGTFLAYAQFND 230
 181 TEVPLIEYSFYSDESLQYPKTRVTPPKAGAVNPTVKFVYNTDSLSVTNATSIOTIAP 240
 231 TEVPLIEYSFYSDESLQYPKTRVTPPKAGAVNPTVKFVYNTDSLSVTNATSIOTIAP 290
 241 ASMLIGDHYLCDVWTATQERISLQMLRRIQNYSMDICDYESSGRNCLVARQHTEMST 300
 291 ASMLIGDHYLCDVWTATQERISLQMLRRIQNYSMDICDYESSGRNCLVARQHTEMST 350
 301 TGDWGRFRPSPBHFITDGNSPYK1ISNEEGYRHICFQIDKDCDFITKGTVWIGIEAL 360
 351 TGDWGRFRPSPBHFITDGNSPYK1ISNEEGYRHICFQIDKDCDFITKGTVWIGIEAL 410
 361 TSDLYYYISNEYKGMPGGRNLKYKIQLSDYTKVTCCLSCELLPERCOYVSFSKEAKYQL 420
 411 TSDLYYYISNEYKGMPGGRNLKYKIQLSDYTKVTCCLSCELLPERCOYVSFSKEAKYQL 470
 421 RCGSGPSPFLYLTHSSYNDKGRLVERLEDNSALDKMLQNVQMPSKKDFILNETKFWQMIL 480
 471 RCGSGPSPFLYLTHSSYNDKGRLVERLEDNSALDKMLQNVQMPSKKDFILNETKFWQMIL 530
 531 PPHFDKSCKPFLLDVYAGPCSQADTVFLNWATYLASTENIIVASPDGSGYQGDKI 540
 591 MHAINRRLGTEPEVDQEAAQFSKMGFVDNKRIALWGMSYGGYTTSMNGSGSVEFKCG 600
 591 MHAINRRLGTEPEVDQEAAQFSKMGFVDNKRIALWGMSYGGYTTSMNGSGSVEFKCG 650
 601 IAVAPYSRWEEYDSUTERMGLPTPEDNDHYRASTVMSRAENPKQVEYLILHGTDNN 660
 651 IAVAPYSRWEEYDSUTERMGLPTPEDNDHYRASTVMSRAENPKQVEYLILHGTDNN 710
 661 VHFQOQAQISKALVGDFOAMMFTDEDGIGIASSTAHQHITYTMSPFIKOCFSLP 716
 711 VHFQOQAQISKALVGDFOAMMFTDEDGIGIASSTAHQHITYTMSPFIKOCFSLP 766

RESULT 12
 US-10-476-264-147
 / Sequence 147, Application US/10476264
 / Publication No. US2005012310A1
 / GENERAL INFORMATION:
 / APPLICANT: Cookson, William Osmond Charles Michael
 / APPLICANT: Moffat, Miriam Fleur
 / APPLICANT: Allen, Maxine
 / APPLICANT: Lench, Nick
 / TITLE OF INVENTION: Enzyme and SNP marker for disease
 / FILE REFERENCE: 16721-002US1
 / CURRENT APPLICATION NUMBER: US/10/476,264
 / CURRENT FILING DATE: 2003-10-24
 / PRIOR APPLICATION NUMBER: PCT/GB02/01887
 / PRIOR APPLICATION NUMBER: GB0110044-5
 / PRIOR APPLICATION NUMBER: GB0110046-0
 / PRIOR FILING DATE: 2001-04-24
 / PRIOR APPLICATION NUMBER: GB0110046-0
 / PRIOR FILING DATE: 2001-04-24
 / PRIOR APPLICATION NUMBER: GB0124594-3
 / PRIOR FILING DATE: 2001-10-12
 / PRIOR APPLICATION NUMBER: GB0124575-2
 / PRIOR FILING DATE: 2001-10-12
 / NUMBER OF SEQ ID NOS: 1
 / SOFTWARE: PatentIn version 3.1
 / NUMBER OF SEQ ID NOS: 1
 / SOFTWARE: PatentIn version 3.1

; SEQ ID NO 147
 ; LENGTH: 766
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-476-264-147
 Query Match 99.7%; Score 3867; DB 5; Length 766;
 Best Local Similarity 99.7%; Bred. No. 3.9e-311;
 Matches 714; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 NTYRLKYSLRWISDHELYKQENNVLVNAEYGNSSVFLNSTDFEGHSINDYSISPD 60
 Db 51 NTYRLKYSLRWISDHELYKQENNVLVNAEYGNSSVFLNSTDFEGHSINDYSISPD 60
 Qy 61 QOFILLENTYKQWRHSTSAYDIYDLNKRLITEERIPNNTOWTWSPVGHKLAYWN 120
 Db 111 QOFILLENTYKQWRHSTSAYDIYDLNKRLITEERIPNNTOWTWSPVGHKLAYWN 120
 Qy 111 QOFILLENTYKQWRHSTSAYDIYDLNKRLITEERIPNNTOWTWSPVGHKLAYWN 170
 Db 121 DIFYKIEPNLPSRITYTGKEDIYNGTIDMWYEEVSAYSALWNSPNTFLAYAQFND 180
 Qy 121 DIFYKIEPNLPSRITYTGKEDIYNGTIDMWYEEVSAYSALWNSPNTFLAYAQFND 180
 Db 171 DIFYKIEPNLPSRITYTGKEDIYNGTIDMWYEEVSAYSALWNSPNTFLAYAQFND 230
 Qy 171 DIFYKIEPNLPSRITYTGKEDIYNGTIDMWYEEVSAYSALWNSPNTFLAYAQFND 230
 Db 181 TEVPLIEYSFSYSDSLOQPKTVTRVPYPKAGAVNPPTVKFVVNTDSLSSVTNATSIQTAP 240
 Qy 181 TEVPLIEYSFSYSDSLOQPKTVTRVPYPKAGAVNPPTVKFVVNTDSLSSVTNATSIQTAP 240
 Db 231 TEVPLIEYSFSYSDSLOQPKTVTRVPYPKAGAVNPPTVKFVVNTDSLSSVTNATSIQTAP 290
 Qy 231 TEVPLIEYSFSYSDSLOQPKTVTRVPYPKAGAVNPPTVKFVVNTDSLSSVTNATSIQTAP 290
 Db 241 ASMLGDHYLCDVTTWATERISLQWLRRIONSYMDICDYEESGRWNCLVARQHTEMST 300
 Qy 241 ASMLGDHYLCDVTTWATERISLQWLRRIONSYMDICDYEESGRWNCLVARQHTEMST 300
 Db 291 ASMLGDHYLCDVTTWATERISLQWLRRIONSYMDICDYEESGRWNCLVARQHTEMST 350
 Qy 301 TCGWGRFRPSEPHFTLQGNSFYKIIISNEGGYHICYFQIDKKDCFTTKGTWVIGTEAL 360
 Db 351 TCGWGRFRPSEPHFTLQGNSFYKIIISNEGGYHICYFQIDKKDCFTTKGTWVIGTEAL 410
 Qy 301 TSDLYLTSNEYKGMPGGRNLXKIQLDYTYKTCSCLNPERCQYSYSFSKEAKYQOL 420
 Db 351 TSDLYLTSNEYKGMPGGRNLXKIQLDYTYKTCSCLNPERCQYSYSFSKEAKYQOL 420
 Qy 361 TSDLYLTSNEYKGMPGGRNLXKIQLDYTYKTCSCLNPERCQYSYSFSKEAKYQOL 420
 Db 411 TSDLYLTSNEYKGMPGGRNLXKIQLDYTYKTCSCLNPERCQYSYSFSKEAKYQOL 470
 Qy 411 TSDLYLTSNEYKGMPGGRNLXKIQLDYTYKTCSCLNPERCQYSYSFSKEAKYQOL 470
 Db 421 RCGSPGLPLTHSYNDKGRLVELENSADKMLQNTQMPKKLDPTILNETKFWQMIL 480
 Qy 421 RCGSPGLPLTHSYNDKGRLVELENSADKMLQNTQMPKKLDPTILNETKFWQMIL 480
 Db 471 RCGSPGLPLTHSYNDKGRLVELENSADKMLQNTQMPKKLDPTILNETKFWQMIL 530
 Qy 471 RCGSPGLPLTHSYNDKGRLVELENSADKMLQNTQMPKKLDPTILNETKFWQMIL 530
 Db 481 PHFDKSKYKPYLLDYYAGPCSKOKADTYFRLNWATYLASTENTIVASFDGRGSSYQGDKI 540
 Qy 481 PHFDKSKYKPYLLDYYAGPCSKOKADTYFRLNWATYLASTENTIVASFDGRGSSYQGDKI 540
 Db 531 PHFDKSKYKPYLLDYYAGPCSKOKADTYFRLNWATYLASTENTIVASFDGRGSSYQGDKI 590
 Qy 531 PHFDKSKYKPYLLDYYAGPCSKOKADTYFRLNWATYLASTENTIVASFDGRGSSYQGDKI 590
 Db 541 MAINIRLGTFEVEDQLEAAROFSKMGFVNKRIAIWGWSSGGYTTSMVLGSGSVFKCG 600
 Qy 541 MAINIRLGTFEVEDQLEAAROFSKMGFVNKRIAIWGWSSGGYTTSMVLGSGSVFKCG 600
 Db 591 MAINIRLGTFEVEDQLEAAROFSKMGFVNKRIAIWGWSSGGYTTSMVLGSGSVFKCG 650
 Qy 591 MAINIRLGTFEVEDQLEAAROFSKMGFVNKRIAIWGWSSGGYTTSMVLGSGSVFKCG 650
 Db 601 IAVAPSRWEYDSVTERMGLPTBDNLHYRNSTMSAENFKQVEYLIGHGTADDN 660
 Qy 601 IAVAPSRWEYDSVTERMGLPTBDNLHYRNSTMSAENFKQVEYLIGHGTADDN 660
 Db 651 IAVAPSRWEYDSVTERMGLPTBDNLHYRNSTMSAENFKQVEYLIGHGTADDN 710
 Qy 651 IAVAPSRWEYDSVTERMGLPTBDNLHYRNSTMSAENFKQVEYLIGHGTADDN 710
 Db 661 VHFQOSAQISKALVALDVGVDFQAMWYTDDEDHGIASSTAHOHITYTHMSHTKQCFSLP 716
 Qy 661 VHFQOSAQISKALVALDVGVDFQAMWYTDDEDHGIASSTAHOHITYTHMSHTKQCFSLP 766
 Db 711 VHFQOSAQISKALVALDVGVDFQAMWYTDDEDHGIASSTAHOHITYTHMSHTKQCFSLP 766

RESULT 14
 US-10-770-712-2
 Sequence 2, Application US/10707012
 Publication No. US20050170331A1
 GENERAL INFORMATION:
 ; APPLICANT: Vojdani, Aristo
 ; TITLE OF INVENTION: IDENTIFICATION OF ETIOLOGY OF AUTISM
 ; FILE REFERENCE: INSCI2_008A
 ; CURRENT APPLICATION NUMBER: US/10/770_712
 ; CURRENT FILING DATE: 2004-02-03
 ; NUMBER OF SEQ ID NOS: 133
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 766
 ; TYPE: PRT
 ; ORGANISM: Rattus norvegicus
 US-10-770-712-1
 Sequence 1, Application US/10707012
 Publication No. US20050170331A1
 GENERAL INFORMATION:
 ; APPLICANT: Vojdani, Aristo
 ; TITLE OF INVENTION: IDENTIFICATION OF ETIOLOGY OF AUTISM
 ; FILE REFERENCE: INSCI2_008A
 ; CURRENT APPLICATION NUMBER: US/10/770_712
 ; CURRENT FILING DATE: 2004-02-03
 ; NUMBER OF SEQ ID NOS: 133
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 766

US-10-770-712-2

Query Match 86.6%; Score 3358.5; DB 5; Length 767;
 Best Local Similarity 84.8%; Pred. No. 5.4e-269;
 Matches 609; Conservative 51; Mismatches 55; Indels 3; Gaps 1;

Qy 1 NTYRLKLYSLRWISDHEYLKQENNTLIVNAYGNSSVLENSTDFEFGHSINDYSISPD 60
 Db 49 NTRFRVKSYSLRWVSDBEYLKQENNLFLNAERGNSIIFLENSTEPIFGSISDVSVPD 108
 Qy 61 GQFILLENYNTKQMRHSYTASYDIYDLNKROLITERIPANTONTWSPYGHKLAYVN 120
 Db 109 RLFVLLLENYNTKQMRHSYTASYDIYDLNKROLITEKIPANTONTWSEGHKLAYVN 168
 Qy 121 DIFYKKEPNLSESYRTWTGEDIINGITDWYEEVFSAYSALIWSPNGTEFLAYAQFND 180
 Db 169 DIFYKKEPHLPSHRITSTGENVIENGINDWYEEBEGAYSAIWLMSPNGTEFLAYAQFND 228
 Qy 181 TEPVPLIEYSFYSDESLQYPKTVRVPYPKAGAVNPFTKFFVNTDSLISVTNATSIQITAP 240
 Db 229 TGPVPLIEYSFYSDESLQYPKTVWIPYPKAGAVNPFTKFFVNTDSLSTTTIMQITAP 288
 Qy 241 ASMLIGDHYLCDVTWATERISLWRIRQNYSYMIDCYDESSGRWNCLVARQHIE MST 300
 Db 289 ASVTTGDHYLCDVAVWSEDRISLWRIRQNYSYMACDYLDDKRTLVNWCPTRREIETSA 348
 Qy 301 TGWVGFRPSPSPHFTLDGSNSPYKILSNEEGYRHICYFQLDKRD -- -CTFPTKGTVWEVIGI 357
 Db 349 TGWCGRFRPAEPHFTSDGSSFYKIVSDKDGYKHICQFDQRKPEQVCTFTKGAMEVII 408
 Qy 358 EALTSDYLYTSNEYKGMPGGRNLKYKQLSDPTYKTCSCELNPERCQYTSVSFSKEARY 417
 Db 409 BALTSWLYMISNEYKMPGGRNLKYQDHTNKKCLSDLPNRCQYTSVSFSKEARY 468
 Qy 418 YQLRCSPGGLPLYTLHSYNDKGHLVMTEDNSALDKMLQNYOMPSSKGLDFTLNETKFWQ 477
 Db 469 YQLCGRGPGLPLYTLHRSTDQKELRLVLTEDNSALDKMLQDYQMPSSKCLDPVLNTRFWQ 528
 Qy 478 MILPPHPDKSKKYPPLDVTAGPCSKQADTVFLNWTIALASTENIIVASFDGGSGYQG 537
 Db 529 MILPPHPDKSKKYPPLDVTAGPCSKOKADAFLRNWTIALASTENIIVASFDGRGSGQG 588
 Qy 538 DKIMHAIRRLGTPEVDQTEAAROFSKNGFVDNKRIAIWGSYCGYTSMVJGSGSGYQG 597
 Db 589 DKIMHAIRKRGTLTEVDQTEAARQPLKMGPEVDNKRIAIWGSYCGYTSMVJGSGSGYQG 648
 Qy 598 KCGIAVAPVSRWEYDSSVTERYMLPPTPBDNLHYRNSTUMSRAENFKQEVEYLHGT A 657
 Db 649 KCGIAVAPVSRWEYDSSVTERYMLPPTPBDNLHYRNSTUMSRAENFKQEVEYLHGT A 708
 Qy 658 DDNTEHQFQSQISKALVDGVDFQAMWYTDDBHGJASTAHQHUYTHMSHFIROCFS L 715
 Db 709 DDNTHFQQSQAQISKALVDAGVDFQAMWYTDDBHGJASTAHQHUYTHMSHFQLOQCFS L 766

RESULT 15
 US-10-165-603-4
 Sequence 4, Application US/10/165603
 Publication No. US20030021792A1
 GENERAL INFORMATION:
 i APPLICANT: Roben, Paul W.
 i TITLE OF INVENTION: TISSUE-SPECIFIC ENDOTHELIAL MEMBRANE
 i FILE REFERENCE: PTTECH 001A
 i CURRENT APPLICATION NUMBER: US/10/165,603
 i PRIORITY FILING DATE: 2002-06-07
 i PRIORITY APPLICATION NUMBER: 60/297,021
 i PRIORITY FILING DATE: 2001-06-08
 i PRIORITY FILING DATE: 60/305,117
 i NUMBER OF SEQ ID NOS: 33
 i SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 4
; LENGTH: 767
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; US-10-165-603-4
Query Match 86.3%; Score 3347.5; DB 4; Length 767;
Best Local Similarity 81.5%; Pred. No. 4.4e-268;
Matches 607; Conservative 52; Mismatches 56; Indels 3; Gaps 1;
; NTYRLKLYSLRWISDHEYLKQENNTLIVNAYGNSSVLENSTDFEFGHSINDYSISPD 60
; 49 NTRFRVKSYSLRWVSDBEYLKQENNLFLNAERGNSIIFLENSTEPIFGSISDVSVPD 108
; 61 GQFILLENYNTKQMRHSYTASYDIYDLNKROLITEKIPANTONTWSEGHKLAYVN 120
; 109 RLFLVLLLENYNTKQMRHSYTASYDIYDLNKROLITEEKIPANTONTWSEGHKLAYVN 168
; 121 DIFYKKEPNLSESYRTWTGEDIINGITDWYEEVFSAYSALIWSPNGTEFLAYAQFND 180
; 169 DIFYKKEPHLPSHRITSTGENVIENGINDWYEEBEGAYSAIWLMSPNGTEFLAYAQFND 228
; 229 TGPVPLIEYSFYSDESLQYPKTVWIPYPKAGAVNPFTKFFVNTDSLSTTTIMQITAP 288
; 241 ASVTTGDHYLCDVAVWSEDRISLWRIRQNYSYMIDCYDESSGRWNCLVARQHIE MST 300
; 289 ASVTTGDHYLCDVAVWSEDRISLWRIRQNYSYMACDYLDDKRTLVNWCPTRREIETSA 348
; 301 TGWVGFRFREPHTSDGSSFYKIVSDKDGYKHICQFDQRKPEQVCTFTKGAMEVII 348
; 349 TGWCGRFRPAEPHFTSDGSSFYKIVSDKDGYKHICQFDQRKPEQVCTFTKGAMEVII 357
; 399 TGPVPLIEYSFYSDESLQYPKTVRVPYPKAGAVNPFTKFFVNTDSLISVTNATSIQITAP 408
; 358 EALTSDYLYTSNEYKGMPGGRNLKYKQLSDPTYKTCSCELNPERCQYTSVSFSKEARY 417
; 409 EALTSDYLYTSNEYKGMPGGRNLKYQDHTNKKCLSDLPNRCQYTSVSFSKEARY 468
; 418 YQLRCSPGGLPLYTLHSYNDKGHLVMTEDNSALDKMLQNYOMPSSKGLDFTLNETKFWQ 477
; 469 YGLGORGPGSLPLYTLHRSTDQKELRLVLTEDNSALDKMLQDVQMPSSKCLDPVLNTRFWQ 528
; 478 MILPPHPDKSKKYPPLDVTAGPCSKQADTVFLNWTIALASTENIIVASFDGGSGYQG 537
; 529 MILPPHPDKSKKYPPLDVTAGPCSKOKADAFLRNWTIALASTENIIVASFDGRGSGQG 588
; 538 DKIMHAIRRLGTPEVDQTEAAROFSKNGFVDNKRIAIWGSYCGYTSMVJGSGSGYQG 597
; 589 DKIMHAIRKRGTLTEVDQTEAARQPLKMGPEVDNKRIAIWGSYCGYTSMVJGSGSGYQG 648
; 649 KCGIAVAPVSRWEYDSSVTERYMLPPTPBDNLHYRNSTUMSRAENFKQEVEYLHGT A 657
; 709 DDNTHFQQSQAQISKALVDAGVDFQAMWYTDDBHGJASTAHQHUYTHMSHFQLOQCFS L 708

Search completed: January 28, 2006, 00:04:55
Job time : 73 secs

THIS PAGE BLANK (USPTO)

Copyright (c) 1993 - 2006	GenCore version 5.1.6	CompuGen Ltd.
Run on:	January 30, 2006, 14:07:15 ; Search time 135 Seconds (without alignments) 2330.335 Million cell updates/sec	
Title:	US-10-659-055-1_COPY_51_766	
Perfect score:	1	NTRKLKSLRWWSDHEVLY.....AHOIYTHMSHFKQCFSLP 716
Sequence:		
Scoring table:	BLOSUM62	
	Gapop 10.0 , Gapext 0.5	
Searched:	2443163 seqs, 439378781 residues	
Total number of hits satisfying chosen parameters:	2443163	

post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Maximum DB seq length: 2000000000

Post-processing: Minimatch 0% Maximum Match 100%
 Listing first 45 summaries

Database : A_Geneset_21.*

RESULT 1
ID AAR54612 standard; protein; 759 AA.
vv.

ALIGNMENTS

1:	Genesegp2008::	AC	AARK51612;
2:	Genesegp1980B::*	XX	
3:	Genesegp2000B::*	DT	25-MAR-2003 (revised)
4:	Genesegp2001B::*	DT	09-DEC-1994 (first entry)
5:	Genesegp2002B::*	XX	
6:	Genesegp2003as::*	DB	Delta3-9 CD26.
7:	Genesegp2003bs::*	XX	
8:	Genesegp2004B::*	KW	Human; T cell activation antigen; CD26; analogues; deletion; soluble;
9:	Genesegp2005B::*	KW	signal peptidase; immune-stimulating; response-stimulating; AIDS;
		KW	immunosuppression; AIDS-related complex.
		XX	
		OS	Homo sapiens.

ALIGNMENTS

ALIGNMENTS

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :	A_Geneseq_21:*	RESULT 1 AAR54612
	1: geneseqp1900s:*	ID AAR54612 standard; protein; 759 AA.
	2: geneseqp1900s:*	XX
	3: geneseqp2000s:*	AC
	4: geneseqp2001s:*	XX
	5: geneseqp2002s:*	25-MAR-2003 (revised)
	6: geneseqp2003bs:*	DT 09-DEC-1994 (first entry)
	7: geneseqp2003bs:*	XX
	8: geneseqp2004s:*	Delta3-9 CD26.
	9: geneseqp2005s:*	XX
		DE
		XX
		Human; T cell activation antigen; CD26; analogues; deletion; soluble;
		KW signal peptidase; immune-stimulating; response-stimulating; AIDS;
		KW immunosuppression; AIDS-related complex.
		XX
		Homo sapiens.
		OS
		XX
		FH Key
		FT Misc-difference 2..3
		FT /note= "Position of delta3-9 deletion"

Pred. No. 18 is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

No.	Score	Match	Length	DB ID	Description	PN
1	3877	100.0	759	2	AAR54612 Delta3-9	WO49132-A1 . XX
2	3877	100.0	766	2	AAR40909 Sequence	PD XX
3	3877	100.0	766	5	AbB08991 Human dip	XX PF
4	3877	100.0	766	5	AAG78417 Human dip	19-AUG-1993; XX
5	3877	100.0	766	7	ADD27855 Human dip	93WO-US007923. PR
6	3877	100.0	766	7	Add46934 Human Pro	21-AUG-1992; XX
7	3877	100.0	766	7	Adn39272 Cancer INST INC.	(DAND) DANA FARBER CANCER INST INC. PA
8	3877	100.0	766	8	Adj83981 Human ful	XX
9	3877	100.0	766	8	Adj75313 Marker ge	P1 Morimoto C, Schlossman S, Tanaka T;
10	3877	100.0	766	8	Ado19398 Human PRO	XX
11	3877	100.0	766	8	Ado19806 Human PRO	WPI : 1994-151317/18. DR
12	3877	100.0	766	8	Ado71612 Amino aci	XX
13	3877	100.0	766	8	Ado71644 Amino aci	PT
14	3877	100.0	766	8	ABM80355 Tumour-as	PT
15	3877	100.0	766	8	Adp54458 Human PRO	PT
16	3877	100.0	766	8	Adu06688 Novel bro	XX
17	3877	100.0	766	8	Adv25525 Human dip	PS Claim 3 : Page 49-52; 85pp; English.
18	3877	100.0	766	9	ADY15161 PRO polyp	XX
19	3877	100.0	766	9	Ady16580 PRO polyp	CC
20	3877	100.0	766	9	Adz14038 Human dip	CC
21	3877	100.0	766	9	Aeb94223 CD26/dipe	CC
22	3871	99.8	736	8	Ado40240 Human DPP	CC
23	3871	99.8	736	8	Abg61910 Prostate	CC
24	3871	99.8	766	5	Aao15555 Human din	CC
					ABG61910.5	CC

eg. they may be used for treatment of disease conditions characterised by immunosuppression, eg. AIDS or AIDS-related complex, other virally or environmentally-induced conditions, and certain congenital immune deficiencies. The peptides can be employed to increase immune function which has been impaired by use of immunosuppressive drugs, such as certain chemotherapeutic drugs. (Updated on 25-MAR-2003 to correct PN field.)

Sequence 759 AA;

Query Match	100.0%	Score 3877;	DB 2;	Length 759;
Best Local Similarity	100.0%	Pred. No. 0;	Gaps 0;	
Matches	716;	Conservative 0;	Mismatches 0;	
Indels	0;			
SQ				

```

QY      1 NTYRLKLYSLRWSIDHELYKQENNIVLVAEYGNSSVLENSTFDEFGHSSINDYSISPD 60
Db      44 NTYRLKLYSLRWSIDHELYKQENNIVLVAEYGNSSVLENSTFDEFGHSSINDYSISPD 103
QY      61 GQFILEENYYVKQWHSYTASYDIDYLNRQLTIEERIPNTQTVTWSPVGHLAYWN 120
Db      104 GQFILEENYYVKQWHSYTASYDIDYLNRQLTIEERIPNTQTVTWSPVGHLAYWN 163
QY      121 DIVYKIEPNLPSYRITWIGKEDITYINGITDWVYBEEVSAYSALWWSPNGTFLAYAQND 180
Db      164 DIVYKIEPNLPSYRITWIGKEDITYINGITDWVIEBEVSAYSALWWSPNGTFLAYAQND 223
QY      181 TEVPLIEYSPSYSDESLQYPKTVRYPKAGAVANPTVKFVNTDLSVTNTNATSIQTAP 240
Db      224 TEVPLIEYSPSYSDESLQYPKTVRYPKAGAVNPTVKFVNTDLSVTNTNATSIQTAP 283
QY      241 ASMLIGDHYLICDVTWATQERISLWLRQIONYSYMIDCYDESSGRWNLVAQHIENT 300
Db      284 ASMLIGDHYLICDVTWATQERISLWLRQIONYSYMIDCYDESSGRWNLVAQHIENT 343
QY      301 TGWGRFRPSEPHFTLGNSFYKITSNEEGSYRHYCQFDKDDCTFPGTWEGITAL 360
Db      344 TGWGRFRPSEPHFTLGNSFYKITSNEEGSYRHYCQFDKDDCTFPGTWEGITAL 403
QY      361 TSDYLIXTSNEYKGMPGGRNLKYKTSYDVKTCUCLSCELNPERCQYYSFSKEAKYKQL 420
Db      404 TSDYLIXTSNEYKGMPGGRNLKYKTSYDVKTCUCLSCELNPERCQYYSFSKEAKYKQL 463
QY      421 RCGSPGLPLYTLHSSYNDKCLRVLEDNSALDKMLQNVQMPSKCLDFILNETKFWQML 480
Db      464 RCGSPGLPLYTLHSSYNDKCLRVLEDNSALDKMLQNVQMPSKCLDFILNETKFWQML 523
QY      481 PPHFDKSCKYPLLDVYAGPCSKQADTVPLRNATYLASTENITVASTFGRGSYQGKRI 540
Db      524 PPHFDKSCKYPLLDVYAGPCSKQADTVPLRNATYLASTENITVASTFGRGSYQGKRI 583
QY      541 MHAIRNRLGTFEVEDQIEARQFSKMGYVDNKRAIKWMSYGGVTFSMVTLGSSEGVFCG 600
Db      584 MHAIRNRLGTFEVEDQIEARQFSKMGYVDNKRAIKWMSYGGVTFSMVTLGSSEGVFCG 643
QY      601 IAVAPVSRVEYYDSVYTERYMLPTPENLDHYRNTSTMRAENFKQVEYLJLHGTAQN 660
Db      644 IAVAPVSRVEYYDSVYTERYMLPTPENLDHYRNTSTMRAENFKQVEYLJLHGTAQN 703
QY      661 VHFQOSA1QSKALVQDVGDFEQAMMYTDDEDHGIASTAHQIYTMHSHFIKQCFSLP 716
Db      704 VHFQOSA1QSKALVQDVGDFEQAMMYTDDEDHGIASTAHQIYTMHSHFIKQCFSLP 759

```

KW	Human T cell activation antigen; monoclonal antibody Tai.
XX	
OS	Homo sapiens
XX	
FH	Key
Region	Location/Qualifiers
FT	7
Region	/label= hydrophobic
FT	29 . 323
FT	/label= N-terminal glycosylated region of extracellular domain
FT	/note= "8 sites for N-linked glycans"
FT	324 . 551
FT	/label= Cysteine rich region of extracellular domain
FT	/note= "1 N-linked glycosylation site"
FT	552 . 766
FT	/label= C-terminal region of extracellular domain
FT	/note= "1 N-linked glycosylation site & 1 catalytic site"
FT	627 . 631
FT	/label= active site of serine protease/esterase
FT	/note= "fits the consensus sequence GXSXG"
XX	
PN	W09316102-A1.
XX	
XX	19-AUG-1993.
XX	09-APR-1992;
XX	92WO-US002892.
XX	06-FEB-1992;
XX	92US-00832211.
XX	(DAND) DANA FARBER CANCER INST INC.
XX	Morimoto C, Schlossman SF, Tanaka T;
XX	WPI; 1993-272827/34.
DR	N-PSDB; AAQ16089.
XX	
PT	Polypeptide fragments of CD26 - are capable of disrupting binding of CD45 and CD26 and thus interfering with T-cell activation.
XX	
XX	Disclosure; Page 39-43; 73pp; English.
XX	
CC	C26 is a human T cell activation antigen originally identified by its reactivity with the Mab Tai. C26 cDNA library was constructed from human PBMC-activated T cells using the CDM vector. The hydrophobic N-terminal of the predicted C26 polypeptide has the characteristics of a signal sequence of the type II membrane protein, which is reinforced by the observation that potential N-glycosylation sites are located in the carboxy side of the hydrophobic core. Therefore the N-terminal 6 AAs are predicted to be cytoplasmic, the next 22 AAs are predicted to transverse the cytoplasmic membrane, and the 738 C-terminal AAs constitute the predicted extracellular domain. (Updated on 25-MAR-2003 to correct PN field.)
CC	
CC	CD26 is a human T cell activation antigen originally identified by its reactivity with the Mab Tai. C26 cDNA library was constructed from human PBMC-activated T cells using the CDM vector. The hydrophobic N-terminal of the predicted C26 polypeptide has the characteristics of a signal sequence of the type II membrane protein, which is reinforced by the observation that potential N-glycosylation sites are located in the carboxy side of the hydrophobic core. Therefore the N-terminal 6 AAs are predicted to be cytoplasmic, the next 22 AAs are predicted to transverse the cytoplasmic membrane, and the 738 C-terminal AAs constitute the predicted extracellular domain. (Updated on 25-MAR-2003 to correct PN field.)
CC	
PS	Sequence 766 AA;
XX	
QY	1 NTYRLKLYSLRWSIDHELYKQENNIVLVAEYGNSSVLENSTFDEFGHSSINDYSISPD 60
Db	51 NTYRLKLYSLRWSIDHELYKQENNIVLVAEYGNSSVLENSTFDEFGHSSINDYSISPD 60
Db	51 NTYRLKLYSLRWSIDHELYKQENNIVLVAEYGNSSVLENSTFDEFGHSSINDYSISPD 110
QY	1 NTYRLKLYSLRWSIDHELYKQENNIVLVAEYGNSSVLENSTFDEFGHSSINDYSISPD 60
Db	51 NTYRLKLYSLRWSIDHELYKQENNIVLVAEYGNSSVLENSTFDEFGHSSINDYSISPD 60
Db	51 NTYRLKLYSLRWSIDHELYKQENNIVLVAEYGNSSVLENSTFDEFGHSSINDYSISPD 110
QY	1 GQFILLENYVQWHRSYTASIDYDNKROLITEERIPNTNTWVTPSPVGHKLAYWN 120
Db	111 GQFILLENYVQWHRSYTASIDYDNKROLITEERIPNTNTWVTPSPVGHKLAYWN 170
Db	121 DIYKIEPNLPSYRITWIGKEDITYINGITDWVYBEEVSAYSALWWSPNGTFLAYAQND 180
Db	171 DIYKIEPNLPSYRITWIGKEDITYINGITDWVYBEEVSAYSALWWSPNGTFLAYAQND 230
QY	181 TEVPLIEYSPSYSDSLOQPKTVRYPYPKAGAVNPTVKEFVNTDSSLSSVTNATSIQTAP 240
XX	

RESULT 2
 AR40909
 ID AAR40909 standard; protein: 766 AA.
 XX
 AC AAR40909;
 DT 25-MAR-2003 (revised)
 DT 05-FEB-1994 (first entry)
 XX
 DE Sequence encoded by human CD26 cDNA.
 XX

Db	231	TEVPLIEFSYPSDESLOYPKTVRVPYPKAGAVNPTVYKFFVNTDSLSVTNATSQITAP	290	CC antiasthmatic, antiallergic and antiinflammatory in its action. The CC peptidase cleaves at Xaa-Pro sequences and is useful for treating a CC patient for mucosal inflammation associated with rhinitis or sinusitis, CC which is the result of allergies or asthma. This sequence represents a human dipeptidyl peptidase IV
Qy	241	ASMLIGHYHLCPTVWTAQERISLQWLRIRQNYSMVDCIDYDESSGRWNCLVARDIEMST	300	XX
Db	291	ASMLIGHYHLCPTVWTAQERISLQWLRIRQNYSMVDCIDYDESSGRWNCLVARDIEMST	350	
Qy	301	TGWVGFRPSEPHFTLDGNSPKYKIIISNEGYRHICFYQIDKKDCTPITKGTMVEIGIAL	360	
Db	351	TGWVGFRPSEPHFTLDGNSPKYKIIISNEGYRHICFYQIDKKDCTPITKGTMVEIGIAL	410	
Qy	361	TSDLYLYTISNEYKGMPGCRNLKYKIQSLDYTYKTCISCELAPERCQYSVSFSKEAKYQY	420	Query Match 100.0%; Score 3877; DB 5; Length 766;
Db	411	TSDLYLYTISNEYKGMPGCRNLKYKIQSLDYTYKTCISCELNPRCQYSVSFSKEAKYQY	470	Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0; Matches 716; conservative 0;
Qy	421	RCSGPGLPLYLTHSSYNDKGILRVLBEINSALDKMLQNQVQMPSKKLDFIILNETKFVYOMIL	480	Qy 1 NTYRKLYSURWISDHEYLXQOENNLVNAEYGNSSVPLENSTDEFGHISINDYSDISPD 60
Db	471	RCSGPGLPLYLTHSSYNDKGILRVLBEINSALDKMLQNQVQMPSKKLDFIILNETKFVYOMIL	530	Db 51 NTYRKLYSURWISDHEYLXQOENNLVNAEYGNSSVPLENSTDEFGHISINDYSDISPD 110
Qy	481	PPHFDPDKSKKPYPLLDIIVYAGPCSQADTVFRLNWATYLASTENIIVASFDRGRGSGYQGDKI	540	Qy 61 GQFILEEYNVKQRHSYTAASYDYLDNKRQLITEERIPNTQNTWSPGHKLYAYWNN 120
Db	531	PPHFDPDKSKKPYPLLDIIVYAGPCSQADTVFRLNWATYLASTENIIVASFDRGRGSGYQGDKI	590	Db 111 GQFILEEYNVKQRHSYTAASYDYLDNKRQLITEERIPNTQNTWSPGHKLYAYWNN 170
Qy	541	MHAINRLGTPVEDQEAAQFSKNGFVDNKRIAIWGMSYGGYTSMVJGSGSVFKCG	600	Qy 121 DIYKLEPNUPLSPTYRTWTGKEDILYNGITDMVYEEFVSAYSAWMSPNGTFLAYAQFND 180
Db	591	MHAINRLGTPVEDQEAAQFSKNGFVDNKRIAIWGMSYGGYTSMVJGSGSVFKCG	650	Db 171 DIYKLEPNUPLSPTYRTWTGKEDILYNGITDMVYEEFVSAYSAWMSPNGTFLAYAQFND 230
Qy	601	IAVAPVSRWEYYDSVTERYNGLPLTPBDNLDHYRNSTVMRAENPKOVELLINTTADDN	660	Qy 181 TEVPLIEFSYPSDSDELOQPKTVRVPYPKAGAVNPTVYKFFVNTDSLSVTNATSQITAP 240
Db	651	IAVAPVSRWEYYDSVTERYNGLPLTPBDNLDHYRNSTVMRAENPKOVELLINTTADDN	710	Db 231 TEVPLIEFSYPSDSDELOQPKTVRVPYPKAGAVNPTVYKFFVNTDSLSVTNATSQITAP 290
Qy	661	VHFQQAQISKAQLDVGVDFQAMWTFDEDIGIASSTAHQIYTMISHFIKQCFSLP	716	Qy 241 ASMLGDHYLCDVTTATQERISLQWLRRLQYNSMIDCIDYDESSGRWNCLVARDIEMST 300
Db	711	VHFQQAQISKAQLDVGVDFQAMWTFDEDIGIASSTAHQIYTMISHFIKQCFSLP	766	Db 291 ASMLGDHYLCDVTTATQERISLQWLRRLQYNSMIDCIDYDESSGRWNCLVARDIEMST 350
Db				Qy 301 TGWGRFRPSEPHFTLDGNSFYKLTISNEGYRHICFYOLDKDCTFITKSTMEVIGIRAL 360
Qy				Db 351 TGWGRFRPSEPHFTLDGNSFYKLTISNEGYRHICFYOLDKDCTFITKSTMEVIGIRAL 410
Db				Qy 361 TSDLYLYTISNEYKGMPGCRNLKYKIQSLDYTYKTCISCELNPERCQYSVSFSKEAKYQY 420
Db				Db 411 TSDLYLYTISNEYKGMPGCRNLKYKIQSLDYTYKTCISCELNPERCQYSVSFSKEAKYQY 470
Qy				Qy 421 RCGSPGLPLYLTHSSVNDKGILRVLBEINSALDKMLQNQVQMPSKKLDFIILNETKFVYOMIL 480
Db				Db 471 RCGSPGLPLYLTHSSVNDKGILRVLBEINSALDKMLQNQVQMPSKKLDFIILNETKFVYOMIL 530
Qy				Qy 481 PPHFDKSKKPYPLLDIIVYAGPCSQADTVFRLNWATYLASTENIIVASFDRGRGSGYQGDKI 540
Db				Db 531 PPHFDKSKKPYPLLDIIVYAGPCSQADTVFRLNWATYLASTENIIVASFDRGRGSGYQGDKI 590
Qy				Qy 541 MHANRRLGTFEVEDQIEAROFSKMGFYDNKRAIWIHGSGSYGGYTSMVJGSGSVFKCG 600
Db				Db 591 MHANRRLGTFEVEDQIEAROFSKMGFYDNKRAIWIHGSGSYGGYTSMVJGSGSVFKCG 650
Qy				Qy 601 IAVAPVSRMEYYDSVTERYNGLPLTPBDNLDHYRNSTVMRAENPKOVELLINTTADDN 660
Db				Db 651 IAVAPVSRMEYYDSVTERYNGLPLTPBDNLDHYRNSTVMRAENPKOVELLINTTADDN 710
DE				Qy 661 VHFQQAQISKAQLDVGVDFQAMWTFDEDIGIASSTAHQIYTMISHFIKQCFSLP
DE				Db 711 VHFQQAQISKAQLDVGVDFQAMWTFDEDIGIASSTAHQIYTMISHFIKQCFSLP
XX				RESULT 3
ID	ABB08991	Human dipeptidyl peptidase IV; antiasthmatic; antiallergic; antiinflammatory.	766 AA.	
XX	ABB08991			
AC	ABB08991;			
XX				
DT	19-JUN-2002	(first entry)		
XX				
DE	Human dipeptidyl peptidase IV.			
XX				
OS	Homo sapiens.			
OS	US6337069-B1.			
PN				
XX	08-JAN-2002.			
XX	28-FEB-2001;	2001US-00794236.		
XX	28-FEB-2001;	2001US-00794236.		
PR	(BMRA-) BMRA CORP BV.			
XX				
P1	Grouzmann E, Lacroix J, Monod M;			
XX	WPI: 2002-16335/21.			
DR				RESULT 4
XX	Treating a patient for mucosal inflammation associated with rhinitis, sinusitis or both, by intranasally administering a peptidase that cleaves at Xaa-Pro sequences, to the patient.			AAG78417
PT				ID AAG78417 standard; protein: 766 AA.
PT				XX AC AAG78417;
PT				XX DT 12-APR-2002 (first entry)
XX	Thus invention relates to the treating of a patient for mucosal inflammation associated with rhinitis or sinusitis, comprising intranasally administering a peptidase. The peptidase is considered			DE Human dipeptidyl peptidase IV amino acid sequence.
CC				XX KW 21953 prolyl oligopeptidase; antibody; proline; endopeptidase; cancer; cardiovascular disease; autoimmune disease; atopc allergy;
CC				CC

KW neuronal disorder; vascular disorder; prostate disorder; cytostatic;
 KW antidiabetic; antiarthritic; antisthmatic; antiflammatory;
 KW multiple sclerosis; asthma;
 KW Grave's disease; neuronal disorder; demyelinating disease;
 KW dipeptidyl peptidase.
 OS Homo sapiens.
 XX WO200179473-A2.
 XX 25-OCT-2001.
 PD XX 11-APR-2001; 2001WO-US040483.
 PF XX 18-APR-2000; 2000US-0197508P.
 PA (MILL-) MILLENNIUM PHARM INC.
 XX P1 Meyers RA, Williamson M;
 XX DR; 2002-034353/04.
 XX New polypeptides 21953, member of human prolyl oligopeptidase family,
 PT which is cytostatic, antidiabetic, antiarthritic, neuroprotective,
 PT useful as diagnostic targets and therapeutic agents for controlling
 PT cancer, lymphoma and leukemia.
 XX Disclosure; Fig 3; 121pp; English.
 XX This invention relates to an isolated 21953 human prolyl oligopeptidase.
 CC Which is cytostatic, antidiabetic, antiarthritic, neuroprotective,
 CC antithyroid, dermatological, antipsoriatic, antiasthmatic,
 CC ophthalmological, antiinflammatory, nootropic, antimarkinsonian,
 CC anticonvulsant, gynaecological, anticoagulant, cardiotonic,
 CC antiatherosclerotic, anorectic, and metabolic or in its action. Uses include
 CC gene therapy, expression or activity of 21953 protein modulator, it is
 CC useful for identifying a compound which binds to it and can be used in
 CC preventing, treating or detecting a cellular proliferative or
 CC differentiative disorder. The 21953 molecules can act as novel diagnostic
 CC targets and therapeutic agents for controlling disorders associated with
 CC the aberrant activity or degradation of peptide hormones e.g., disorders
 CC associated with cell differentiation and proliferation such as cancer.
 CC immune function, reproductive, neurological and cardiovascular function.
 CC The 21953 molecules are thus useful for treating and preventing cellular
 CC proliferative and differentiative disorders, hematopoietic neoplastic
 CC disorders, immune disorders such as autoimmune diseases, diabetes
 CC mellitus, arthritis, multiple sclerosis, asthma, Grave's disease,
 CC neuronal disorders, demyelinating diseases, vascular disorders and
 CC metabolic or pain disorders. This sequence represents the amino acid
 CC sequence of human dipeptidyl peptidase IV
 XX Sequence 766 AA;

Query Match 100.0%; Score 3877; DB 5; Length 766;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 04-SEP-2003.
 XX 27-NOV-2001; 2001US-00993959.

XX 28-FEB-2001; 2001US-00794236.

XX (BMRA-) BMRA CORP BV.

XX PI Grouzmann E, Lacroix J, Monod M;

XX DR WPI; 2003-811386/76.

XX Treatment of patient for mucosal inflammation associated with rhinitis
 PT and/or sinusitis involves intranasally administering peptidase that
 PT cleaves at Xaa-Pro sequences or agent inhibiting binding of Sp to
 PT neurokinin 1 receptor.

XX Disclosure; SEQ ID NO 1; 14PP; English.

XX The present invention relates to a method of treating a patient for
 CC mucosal inflammation associated with rhinitis and/or sinusitis. The
 CC method comprises intranasally administering to the patient a peptidase

XX 291 ASMLIGDHYLCDVTATQERISLQWLRIQNYSTMIDCYDESSGRNCLVARQHIENST 350
 XX 301 TGWYGRFRPSEPHFTLDGNSFYKLIISNEGYRHICYFOIDKKDCTFIFTKGTVWEVIGEAL 360
 XX 351 TGWYGRFRPSEPHFTLDGNSFYKLIISNEGYRHICYFQIDDKDCTFIFTKGTVWEVIGEAL 410
 XX 361 TSDLYXYSNEYKGMPGGRNLKQLSPDYTKTCLSCPLNPERCQYISVSFSKPKAKYQYL 420
 XX 411 TSDLYXYSNEYKGMPGGRNLKQLSPDYTKTCLSCPLNPERCQYISVSFSKPKAKYQYL 470
 XX 421 RCSCPGLPLYTHLSVNNDKGLRVEEDNSALDKMLQNYQMPSKKLLDFILINETKFWQYQML 480
 XX 471 RCSCPGLPLYTHLSVNNDKGLRVEEDNSALDKMLQNYQMPSKKLLDFILINETKFWQYQML 530
 XX 481 PPHPFDKSCKKXPPLIILDVTAQPCSKADTVFLRNWATYLASTENITIVASFDGRGSYQGDKI 540
 XX 531 PPHPFDKSCKKXPPLIILDVTAQPCSKADTVFLRNWATYLASTENITIVASFDGRGSYQGDKI 550
 XX 541 MHAINRRIGTFFEVDDQEAAEORFSKMGFVDNKRAIAINGWSYGGYTSHVLGSGSGVFKCG 600
 XX 591 MHAINRRIGTFFEVDDQEAAEORFSKMGFVDNKRAIAINGWSYGGYTSHVLGSGSGVFKCG 650
 XX 601 IAVAVPSWEEYTDSSVTERYMLGPTEPDNLHYRNSTMSRAENPKQVBYLLHTGADDN 660
 XX 651 IAVAVPSWEEYTDSSVTERYMLGPTEPDNLHYRNSTMSRAENPKQVBYLLHTGADDN 710
 XX 661 VHFFQSAQISKALVDPYDGFQAMWYTDDEDHGIASSTAHQHITYTHMSHTIKQCFSLP 716
 XX 711 VHFFQSAQISKALVDPYDGFQAMWYTDDEDHGIASSTAHQHITYTHMSHTIKQCFSLP 766
 RESULT 5
 ADD27855 standard; protein: 766 AA.
 ID ADD27855
 XX ADD27855;
 AC DT 15-JAN-2004 (first entry)
 XX Human dipeptidyl peptidase IV (DPPIV).
 DE XX
 KW XX Muosal inflammation; rhinitis; sinusitis; exopeptidase; substance P; SP;
 KW XX neurokinin 1 receptor; NKL receptor; allergy; asthma; antiallergic;
 KW XX antiinflammatory; antiasthmatic; human; dipeptidyl peptidase IV; DPPIV;
 KW XX enzyme.
 OS XX Homo sapiens.
 XX PN US2003165489-A1.
 PN XX
 XX 04-SEP-2003.
 XX 27-NOV-2001; 2001US-00993959.

XX (BMRA-) BMRA CORP BV.

XX Grouzmann E, Lacroix J, Monod M;

XX DR WPI; 2003-811386/76.

XX Treatment of patient for mucosal inflammation associated with rhinitis
 PT and/or sinusitis involves intranasally administering peptidase that
 PT cleaves at Xaa-Pro sequences or agent inhibiting binding of Sp to
 PT neurokinin 1 receptor.

XX Disclosure; SEQ ID NO 1; 14PP; English.

XX The present invention relates to a method of treating a patient for
 CC mucosal inflammation associated with rhinitis and/or sinusitis. The
 CC method comprises intranasally administering to the patient a peptidase

XX 61 QGPILLEENYVKQWHSYASYDYLNLKQLTBEERPNNTQWTWSPVGHKLAYWN 120
 DB 51 NTYRKLKLSRWSHDEVLYKOENNLVNAEYGNSSYLENSTDFEGRHSINDYNSIPD 60
 XX 1 NTYRKLKLSRWSHDEVLYKOENNLVNAEYGNSSYLENSTDFEGRHSINDYNSIPD 60
 XX 51 NTYRKLKLSRWSHDEVLYKOENNLVNAEYGNSSYLENSTDFEGRHSINDYNSIPD 110
 XX 61 QGPILLEENYVKQWHSYASYDYLNLKQLTBEERPNNTQWTWSPVGHKLAYWN 170
 DB 111 QGPILLEENYVKQWHSYASYDYLNLKQLTBEERPNNTQWTWSPVGHKLAYWN 170
 XX 121 DYYKIEPNLPSYRITWTKGEDIYNGITDWWYBEEVFSAYSALWWSNGTFLAYAQND 180
 DB 171 DYYKIEPNLPSYRITWTKGEDIYNGITDWWYBEEVFSAYSALWWSNGTFLAYAQND 230
 XX 181 TEVPLJEYSPYSDSDELQPKTVRYPKAGAVNPYKPVWNTDSISVVTNATSQIQTAP 240
 DB 231 TEVPLJEYSPYSDSDELQPKTVRYPKAGAVNPYKPVWNTDSISVVTNATSQIQTAP 290
 XX 241 ASMLIGDHYLCDVTATQERISLQWLRIQNYSTMIDCYDESSGRNCLVARQHIENST 300

XX

that cleaves at Xaa-Pro sequences or an agent that inhibits the binding of substance P (SP) to the neurokinin 1 (NKL) receptor. The peptidase is an exopeptidase, preferably selected from human dipeptidyl peptidase IV (DPPIV), human quiescent cell proline dipeptidase, human dipeptidyl peptidase 8, or human attractant. The method is useful for treating a patient for mucosal inflammation associated with rhinitis and/or sinusitis which are the result of allergies or asthma. The invention provides an effective treatment of the inflammation associated with both rhinitis and sinusitis. The present sequence represents human DPPIV.

XX Sequence 766 AA;

```

Query Match      100.0%;  Score 3877;  DB 7;  Length 766;
Best Local Similarity 100.0%;  Pred. No. 0;
Matches 716;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy   1 NTYRLKLYSLRWSDHELYKQENNVLVNAEYGNSVFLENSTDFEFGHSINDYSISPD 60
Db   51 NTYRLKLYSLRWSDHELYKQENNVLVNAEYGNSVFLENSTDFEFGHSINDYSISPD 110
Qy   61 GOFILLEYNYKQWRHSYTASYDIYDNLNKQLITEERIPNTQWTTWSPVGHKLAYVN 120
Db   111 GOFILLEYNYKQWRHSYTASYDIYDNLNKQLITEERIPNTQWTTWSPVGHKLAYVN 170
Qy   121 DIVVKIEPNLPSYRTWTKGEDIYNGITDNYVEEEFSAYSALWNSPNTFLAYAQEND 180
Db   171 DIVVKIEPNLPSYRTWTKGEDIYNGITDNYVEEEFSAYSALWNSPNTFLAYAQEND 230
Qy   181 TEVPLIYFSYFSDESLQPKTVRVPKAGAVNPYTKPFVNTDSLSVTNATSIQTAP 240
Db   231 TEVPLIYFSYFSDESLQPKTVRVPKAGAVNPYTKPFVNTDSLSVTNATSIQTAP 290
Qy   241 ASMLIGHYHLCDVTWATQERISLQWLRIQNYSYMDICDDESSGRWNLCLVARHIEMST 300
Db   291 ASMLIGHYHLCDVTWATQERISLQWLRIQNYSYMDICDDESSGRWNLCLVARHIEMST 350
Qy   301 TGWVGRFRPSSEPHFTLDGNSPKYKIIISNEEGYRHICYFQIDKKDCPTKTCWTWIGIEAL 360
Db   351 TGWVGRFRPSSEPHFTLDGNSPKYKIIISNEEGYRHICYFQIDKKDCPTKTCWTWIGIEAL 410
Qy   361 TSDLYLYISNEYKGMPGGRNYKIQSLDYKTVCLSCLENPERCQYSVFSKEAKYQYL 420
Db   411 TSDLYLYISNEYKGMPGGRNYKIQSLDYKTVCLSCLENPERCQYSVFSKEAKYQYL 470
Qy   421 RCGSPGPPLTYLHSYSTNDKGHLRVLEDNSALDKMLQNQOMPSSKKLDFIILNETKFWYQML 480
Db   471 RCGSPGPPLTYLHSYSTNDKGHLRVLEDNSALDKMLQNQOMPSSKKLDFIILNETKFWYQML 530
Qy   481 PPHEFDKSKYKPYLLDVAAGPCSQKADTVFRLNWATYLASENITIVASFDGRGSYQGDKI 540
Db   531 PPHEFDKSKYKPYLLDVAAGPCSQKADTVFRLNWATYLASENITIVASFDGRGSYQGDKI 590
Qy   541 MHAINRLGTFEVEDOIEARQFSKNGFDVNKRIAIWGSYGGYTSMYLJGSGSVFKCG 600
Db   591 MHAINRLGTFEVEDOIEARQFSKNGFDVNKRIAIWGSYGGYTSMYLJGSGSVFKCG 650
Qy   601 IAVAPVSRWEYDSYTERVNGLPTEDDNLDHYRNSTMVSRAENFKQVEYLTHCTADDN 660
Db   651 IAVAPVSRWEYDSYTERVNGLPTEDDNLDHYRNSTMVSRAENFKQVEYLTHCTADDN 710
Qy   661 VHFQQAQISKAVALDVGVDFQAMWYTEDDGIASSTAHQHLYTMHSFIKQCFSLP 716
Db   711 VHFQQAQISKAVALDVGVDFQAMWYTEDDGIASSTAHQHLYTMHSFIKQCFSLP 766

```

```

Query Match      100.0%;  Score 3877;  DB 7;  Length 766;
Best Local Similarity 100.0%;  Pred. No. 0;
Matches 716;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
Qy   1 NTYRLKLYSLRWSDHELYKQENNVLVNAEYGNSVFLENSTDFEFGHSINDYSISPD 60
Db   51 NTYRLKLYSLRWSDHELYKQENNVLVNAEYGNSVFLENSTDFEFGHSINDYSISPD 110
Qy   61 GOFILLEYNYKQWRHSYTASYDIYDNLNKQLITEERIPNTQWTTWSPVGHKLAYVN 120

```

XX Human Protein AAA52308, SEQ ID NO 12620.
CC Human; pain; neuronal tissue; gene therapy;
CC spinal segmental nerve injury; chronic constriction injury; CCI;
CC spared nerve injury; SNI; Chung.
CC Homo sapiens.
OS Unidentified.
XX WO2003016475-A2.
XX PD 27-FEB-2003.
XX PF 14-AUG-2002; 2002WO-US025765.
XX PR 14-AUG-2001; 2001US-0312147P.
XX PR 01-NOV-2001; 2001US-0343382P.
XX PR 26-NOV-2001; 2001US-0333347P.
XX PA (GEHO) GEN HOSPITAL CORP.
PA (PARB) BAYER AG.
XX WOOLF C, D'URSO D, BEFORTE K, COSTIGAN M;
XX WPI; Page: 1017pp; English.
DR WPI; 2003-268112/26.
DR GENBANK; AAA52308.
XX New composition comprising two or more isolated polypeptides, useful for preparing a medicament for treating pain in an animal.
XX PT preparing a medicament for treating pain in an animal.
XX PS Example 1; Page: 1017pp; English.
XX The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. Gene therapy). The sequence presented is a human protein (described in Table 3 of the specification) which is differentially expressed during pain. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form directly from WIPO at ftp://wipo.int/pub/published_pct_sequences.
XX Sequence 766 AA;

```

Query Match      100.0%;  Score 3877;  DB 7;  Length 766;
Best Local Similarity 100.0%;  Pred. No. 0;
Matches 716;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
Qy   1 NTYRLKLYSLRWSDHELYKQENNVLVNAEYGNSVFLENSTDFEFGHSINDYSISPD 60
Db   51 NTYRLKLYSLRWSDHELYKQENNVLVNAEYGNSVFLENSTDFEFGHSINDYSISPD 110
Qy   61 GOFILLEYNYKQWRHSYTASYDIYDNLNKQLITEERIPNTQWTTWSPVGHKLAYVN 120

```

Db	111	GQFILLEVNVTQWRHSYTASYDYLNLNRQLITEERIPNNTQWVTPVGHKLAYWN	170
Qy	121	DIYVKEIPELNPLSPYRITWTGKEDIDLYNGITDWWYEEVEVSAYSAIWWSPNGTFLAYAQFND	180
Db	171	DIYVKEIPELNPLSPYRITWTGKEDIDLYNGITDWWYEEVEVSAYSAIWWSPNGTFLAYAQFND	230
Qy	181	TEVPLIEYSFSYSDS1QYPKTVRPPYKAGAVNFTVKFFVNTDSLSVSYTNATSIQITAP	240
Db	231	TEVPLIEYSFSYSDS1QYPKTVRPPYKAGAVNFTVKFFVNTDSLSVSYTNATSIQITAP	290
Qy	241	ASMLIGDHDYLCDVDTWATQERISL0MLRRI0NYSYMDICDYEDESSGRWNLVYARQHIESTM	300
Db	291	ASMLIGDHDYLCDVDTWATQERISL0MLRRI0NYSMDCDYEDESSGRWNLVYARQHIESTM	350
Qy	301	TGWGRFRPSEPHFTLDGNSFYKL1SNEECYRHICYFQIDKKDCTPITKGTMEVIGIEAL	360
Db	351	TGWGRFRPSEPHFTLDGNSFYKL1SNEECYRHICYFQIDKKDCTPITKGTMEVIGIEAL	410
Qy	361	TSDLYY1SNEYKGMGGRRNLKYQLSDTYTKVTCISCEINPERQOYVSFSKAEKTYQL	420
Db	411	TSDLYY1SNEYKGMGGRRNLKYQLSDTYTKVTCISCEINPERQOYVSFSKAEKTYQL	470
Qy	421	RCSGPGLPLTLLHSSYNDKGRLVLEDNSAALDKM1QNVOMPSKCLDFILNETKFWYQML	480
Db	471	RCSGPGLPLTLLHSSNDKGRLVLEDNSAALDKM1QNVOMPSKCLDFILNETKFWYQML	530
Qy	481	PPHFDKSCKYKPLLDYYAGPCSQAKADTVPLNATYLASTEN1IVASPDGRGSGYQGDK1	540
Db	531	PPHFDKSCKYKPLLDYYAGPCSQAKADTVPLNATYLASTEN1IVASPDGRGSGYQGDK1	590
Qy	541	MHA1NRLGTFEVEDOIEARQFSKMGFDVNKR1AIWGMYSGGVTVTSMLGSSSGGVFKCG	600
Db	591	MHA1NRLGTFEVEDOIEARQFSKMGFDVNKR1AIWGMYSGGVTVTSMLGSSSGGVFKCG	650
Qy	601	IAVAPYRSRWEYDSSYTERYMLPTPEDNLDHYRNSTMRAENFKOVEYLLIHTGADDN	660
Db	651	IAVAPYRSRWEYDSSYTERYMLPTPEDNLDHYRNSTMRAENFKOVEYLLIHTGADDN	710
Qy	661	VHFQQSAQISKALDVGVDFQAMMTDEDHGIASSTAHQHITYTMSHFT1KOCFLP	716
Db	711	VHFQQSAQISKALDVGVDFQAMMTDEDHGIASSTAHQHITYTMSHFT1KOCFLP	766
RESULT 7			
ADN39272			
ID	ADN39272	standard; protein; 766 AA.	
XX	AC	ADN39272;	
XX	DT	17-JUN-2004 (first entry)	
XX	DB	Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:590.	
XX	XX	Human; differential expression; cancer; angiogenic disorder; fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis; inflammatory disease; autoimmune disease; retinal neovascularisation syndrome; scarring; uterine fibroid; detection; diagnosis; prognosis; drug screening; drug targeting; wound healing; contraception; cytostatic; cardiant; immunomodulatory; pulmonary; gene therapy; vaccine.	
XX	XX	Homo sapiens.	
OS	XX	WO2003042651-A2.	
PD	XX	22-MAY-2003.	
PF	XX	13-NOV-2002; 2002WO-US036810.	
PR	XX	13-NOV-2001; 2001US-0350666P.	
PR	XX	21-NOV-2001; 2001US-033264P.	
PR	XX	29-NOV-2001; 2001US-033493P.	
PR	XX	03-DEC-2001; 2001US-033534P.	

Oy 301 TGWVGRFRPSBEPHTLDGNSFYKIIISNEEGYRHYCIFYQIDKKDCPFITKGQTWEVIGEAL 360
 Db 351 TGWVGRFRPSBEPHTLDGNSFYKIIISNEEGYRHYCIFYQIDKKDCPFITKGQTWEVIGEAL 410
 Oy 361 TSDLYIXSNEYKGMPGRNLYKIQSLDPTKTYCLSCELNPERCQYSYSFSKEAKYQL 420
 Db 411 TSDLYIXSNEYKGMPGRNLYKIQSLDPTKTYCLSCELNPERCQYSYSFSKEAKYQL 470
 Oy 421 RCGSPGPPLYLTHSSYNDKGRLVRLEDNSALDKMLQNQMPSKKLDFIILNETKEWYQML 480
 Db 471 RCGSPGPPLYLTHSSYNDKGRLVRLEDNSALDKMLQNQMPSKKLDFIILNETKEWYQML 530
 Qy 481 PPHFDKSKKCPPLLDDYAGPCSQADTVFRLNWTAYLASTENITIVASFDRGSGYQGDKI 540
 Db 531 PPHFDKSKCPCPLLDDYAGPCSQADTVFRLNWTAYLASTENITIVASFDRGSGYQGDKI 590
 Qy 541 MHAINRRLGTFFEDQEAAQRQFSRKMFDVNKRATIAGWSYGGYTSMVLGSGSGVPGCG 600
 Db 591 MHAINRRLGTFFEDQEAAQRQFSRKMFDVNKRATIAGWSYGGYTSMVLGSGSGVPGCG 650
 Qy 601 IAVAPSRWEYDSVTTERMGLPPEDIGIASSTAHQIYTMHSFKQCFSLP 716
 Db 651 IAVAPSRWEYDSVTTERMGLPPEDIGIASSTAHQIYTMHSFKQCFSLP 710
 Qy 661 VHFQOSAQISKALVDGVDFQAMWTTDEDIGIASSTAHQIYTMHSFKQCFSLP 766
 Db 711 VHFQOSAQISKALVDGVDFQAMWTTDEDIGIASSTAHQIYTMHSFKQCFSLP 765

RESULT 8
 ADJ83981 Human full-length colon dipeptidyl peptidase IV (DPPIV) protein.
 ID ADJ83981 standard; protein; 766 AA.
 XX ADJ83981;
 AC :
 DT 06-MAY-2004 (first entry)
 DE Human full-length colon dipeptidyl peptidase IV (DPPIV) protein.
 KW crystal; protein co-ordinate data; dipeptidyl peptidase IV; DPPIV;
 KW immunomodulatory; anti-diabetic; anti-inflammatory; neuroprotective;
 KW antithyroid; antirheumatic; antiarthritic; anti-HIV; cytostatic;
 KW immune response; diabetes; inflammation; multiple sclerosis;
 KW Grave's disease; chronic rheumatoid arthritis; AIDS; cancer; human;
 KW colon; enzyme.
 XX Homo sapiens.
 XX WO2004011640-A1.
 PD 05-FEB-2004.
 XX 28-JUL-2003; 2003WO-JP009523.
 PR 29-JUL-2002; 2002US-0398761P.
 XX (TANA) TANABE SEIYAKU CO.
 XX Hiramatsu H, Kyono K, Shima H, Sugiyama S;
 XX WPI; 2004-156630/15.
 DR N-PSDB; ADJ83980.

XX New crystal of dipeptidyl peptidase IV capable of analyzing its three-dimensional structure, useful for designing, identifying, evaluating or searching an effector of the dipeptidyl peptidase IV.
 PS Claim 3; SEQ ID NO 2; 332pp; English.
 XX The invention relates to a novel crystal of a dipeptidyl peptidase IV (DPPIV) which is sufficient to ensure a resolution capable of analysing its three-dimensional structure up to the side chain level by X-ray crystallographic structural analysis. The crystal of the invention

CC demonstrates immunomodulatory, antidiabetic, antiinflammatory,
 CC neuroprotective, antithyroid, antiarthritic, anti-HIV and
 CC cytosstatic activities and may be useful for providing a three-dimensional
 CC structural coordinate as the information for designing, identifying,
 CC evaluating or searching for an effector of the dipeptidyl peptidase IV.
 CC The effector may be useful as a modulatory agent of immune response and
 CC as a therapeutic or prophylactic agent for diabetes, inflammation,
 CC multiple sclerosis, Grave's disease, chronic rheumatoid arthritis, AIDS
 CC or cancer. The current sequence is that of the human full-length colon
 CC dipeptidyl peptidase IV (DPPIV) protein of the invention.
 XX SQ Sequence 766 AA;
 Query Match Score 100.0%; Score 3877; DB 8; Length 766;
 Best local similarity 100.0%; Prod. No. 0;
 Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 NTYRKLYSLRWSIDHEYLQKQENNLVFNNAEYGNSSVPLENSTDEFGRHSINNDYSISPD 60
 Db 51 NTYRKLYSLRWSIDHEYLQKQENNLVFNNAEYGNSSVPLENSTDEFGRHSINNDYSISPD 110
 Qy 61 GQFILLENNYVKQRHHSITASYDYLNRQLITEERIPNNNTQVTVSPVGHKLAYWNN 120
 Db 111 GQFILLENNYVKQRHHSITASYDYLNRQLITEERIPNNNTQVTVSPVGHKLAYWNN 170
 Qy 121 DIYKIEBNPLPSITWTGKEDIYINGITDWWYEEBEVSPAYSALWSPNTFLAYAQFND 180
 Db 171 DIYKIEBNPLPSITWTGKEDIYINGITDWWYEEBEVSPAYSALWSPNTFLAYAQFND 230
 Qy 181 TEVPLIEYSFSYDSLOQPKTVRVPYKAGAVNFTVKEFVVNTDSLSSYNTNATSIQITAP 240
 Db 231 TEVPLIEYSFSYDSLSQIPKTVRVPYKAGAVNFTVKEFVVNTDSLSSYNTNATSIQITAP 290
 Qy 241 ASMIIJDAYLCDVWATQBRISLQWLRLQNYSVMDICDYDESSGRNCLVARGHIEINST 300
 Db 291 ASMIIJDAYLCDVWATQBRISLQWLRLQNYSVMDICDYDESSGRNCLVARGHIEINST 350
 Qy 301 TGWGRFRPSEPHFTLDGNSFYKIIISNEEGYRHYCYPQIDKDDTPIKGTVWEGIEAL 360
 Db 351 TGWGRFRPSEPHFTLDGNSFYKIIISNEEGYRHYCYPQIDKDDTPIKGTVWEGIEAL 410
 Qy 361 TSDLYIISNEYKGMPGRNLYKQLSDYTKTCLSCELNPERCQYSVSFSKEAKYQL 420
 Db 411 TSDLYIISNEYKGMPGRNLYKQLSDYTKTCLSCELNPERCQYSVSFSKEAKYQL 470
 Qy 421 RCSPGGLPLYTLISSYNDKGLRVEDNSALDKMLQNVOMPSPSKLDFIILNETKFWYQML 480
 Db 471 RCSPGGLPLYTLISSYNDKGLRVEDNSALDKMLQNVOMPSPSKLDFIILNETKFWYQML 530
 Qy 481 PPHDKSKYKPLLLDVAGPCSKQADTVFRLNWTAYLASTENITIVASPDGRGSGYQGDKI 540
 Db 531 PPHDKSKYKPLLLDVAGPCSKQADTVFRLNWTAYLASTENITIVASPDGRGSGYQGDKI 590
 Qy 541 MHAINRRLGTFEVEDQTEAAROFSKMGFVDNKRATIANGWSGYGVTSMVLGSGSGVFKCG 600
 Db 591 MHAINRRLGTFEVEDQTEAAROFSKMGFVDNKRATIANGWSGYGVTSMVLGSGSGVFKCG 650
 Qy 601 IAVAPVSEWEYDSVYTERMGLPTPEINLDAHRSNTMWSRAENFKQVEYLILHGTAADDN 660
 Db 651 IAVAPVSEWEYDSVYTERMGLPTPEINLDAHRSNTMWSRAENFKQVEYLILHGTAADDN 710
 Qy 661 VHQQSAQTSKALWDGYDFQAMMYTDHDHGIASSTAHQIHTTHMSHPIKQCFSLP 716
 Db 711 VHQQSAQTSKALWDGYDFQAMMYTDHDHGIASSTAHQIHTTHMSHPIKQCFSLP 766

RESULT 9
 ADD75313
 ID ADD75313 standard; protein; 766 AA.
 AC :
 DT 20-MAY-2004 (first entry)
 DE The invention relates to a novel crystal of a dipeptidyl peptidase IV (DPPIV) which is sufficient to ensure a resolution capable of analysing its three-dimensional structure, useful for designing, identifying, evaluating or searching an effector of the dipeptidyl peptidase IV.
 XX

XX	Marker gene related amino acid sequence SEQ ID NO:565.	Db	111 GQFILLENYVKWQRHSYTASYDIDYDNLNRQLITEERI PNNTOWTVTNSPGHKLAYWN 170
DE	bronchial asthma; chronic obstructive pulmonary disease;	Qy	121 DIVKIEENLPSKRLTWGKEDLIVINGTDWVTEEEVFSAYSALWWNSNGTFLAYAQEND 180
KW	respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;	Db	171 DIVKIEENLPSKRLTWGKEDLIVINGTDWVTEEEVFSAYSALWWNSNGTFLAYAQEND 230
KW	gene therapy; marker.	Qy	181 TEVPLIEYSFSYSDSLOVPKTRVPYKAGAVNPTVKEFVNTDSLSEVTNATS IQITAP 240
XX	Homo sapiens.	Db	231 TEVPLIEYSFSYSDSLOVPKTRVPYKAGAVNPTVKEFVNTDSLSEVTNATS IQITAP 290
OS		Qy	241 ASMLIGDHYLCDVWATOBRISSLQMLRRIQNSYMDICDYDESSGRNCLVRQHIENST 300
XX	EP1394274-A2.	Db	291 ASMLIGDHYLCDVWATOBRISSLQMLRRIQNSYMDICDYDESSGRNCLVRQHIENST 350
XX	03-MAR-2004.	Qy	301 TGMYGRPDPSEPHFTLDNSFVKISNEEGYRHICQIDKDCTFTIKGTNEVGIBAL 360
PD	04-AUG-2003; 2003EP-00254857.	Db	351 TGMYGRPDPSEPHFTLDNSFVKISNEEGYRHICYQIDKDCTFTIKGTNEVGIBAL 410
PR	06-AUG-2002; 2002JP-00229312.	Qy	361 TSDLYYTISNEYKMPGGRNLYXKQLSDYTKVTCSELCPNPERCQYVSFSKAKYQYL 420
PA	20-MAR-2003; 2003JP-00077212.	Db	411 TSDLYYTISNEYKMPGGRNLYXKQLSDYTKVTCSELCPNPERCQYVSFSKAKYQYL 470
XX	(GENO-) GENOX RES INC.	Qy	421 RCGPGPLPYLTHSYNDKGRLVLEDNSALDKMLQNYOMPSPSKLDFILNETKFWYOMIL 480
XX	Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K,	Db	471 RCGPGPLPYLTHSYNDKGRLVLEDNSALDKMLQNYOMPSPSKLDFILNETKFWYOMIL 530
DR	WPI: 2004-193155/19.	Qy	481 PPHFDKSKKPKLILDVAGPCSKQADTVFLNWATYLASTENIVASFDGRSSGQSDKI 540
XX	Testing for bronchial asthma or chronic obstructive pulmonary disease by comparing the expression level of a marker gene in a biological sample from a subject with the expression level of the gene in a sample from a healthy subject.	Db	531 PPHFDKSKKPKLILDVAGPCSKQADTVFLNWATYLASTENIVASFDGRSSGQSDKI 590
PT	Example 11; SEQ ID NO 565; 241pp; English.	Qy	541 MHAIRNLRTGFVEBDQEAAEAFQSMGFDVNKRIAITGMWSGYGVTSNVLGSGS GVFKCG 600
PT	The present invention describes a method of testing for bronchial asthma or chronic obstructive pulmonary disease. The method comprises determining the expression level of a marker gene in a biological sample from a subject, comparing the expression level determined with the expression level of the marker gene in a biological sample from a healthy subject, and judging whether the subject has bronchial asthma or chronic obstructive pulmonary disease. The marker gene comprises: (a) a group of genes (S1) whose expression levels increase when respiratory epithelial cells are stimulated with interleukin-13; or (b) a group of genes (S2) whose expression levels decrease when described: (1) a reagent (I) for testing for bronchial asthma or chronic obstructive pulmonary disease; (2) a kit for screening for a candidate compound for a therapeutic agent to treat bronchial asthma or chronic obstructive pulmonary disease; (3) an animal model for bronchial asthma or chronic obstructive pulmonary disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a method for producing an animal model for bronchial asthma or chronic obstructive pulmonary disease; (6) a therapeutic agent for bronchial asthma or chronic obstructive pulmonary disease, comprising the compound, a marker gene or an antisense nucleic acid corresponding to a portion of the marker gene, a ribozyme, a polynucleotide that suppresses the expression of the gene through an RNAi effect or an antibody recognising a protein encoded by a marker gene; and (7) a DNA chip for testing for bronchial asthma or a chronic obstructive pulmonary disease, on which a probe has been immobilised to assay a marker gene. (I) has respiratory and antiasthmatic activities, and can be used in gene therapy. The method is useful for testing for or screening for a therapeutic agent for bronchial asthma or chronic obstructive pulmonary disease. The present sequence is used in the exemplification of the present invention.	Db	591 MDAIRNLRTGFVEBDQEAAEAFQSMGFDVNKRIAITGMWSGYGVTSNVLGSGS GVFKCG 650
PS	XX	Qy	601 IAVAPVSRWEYDSVTERYMGIPTPDNLDHYRNSTMRAENFKQVBYLLIHTGTADDN 660
PS	XX	Db	651 IAVAPVSRWEYDSVTERYMGIPTPDNLDHYRNSTMRAENFKQVBYLLIHTGTADDN 710
PS	XX	Qy	661 VHFQSAQISKALVDVQDFQAMWYTDHDGJASTAHAQHITHMSHPIKOCFSLP 716
PS	XX	Db	711 VHFQSAQISKALVDVQDFQAMWYTDHDGJASTAHAQHITYTHMSHPIKOCFSLP 766
RESULT 10			
		ID AD019398 standard; protein; 766 AA.	
		XX XX	
		AC ADO19398;	
		XX XX	
		DT 12-AUG-2004 (first entry)	
		XX XX	
		DE Human PRO polypeptide #164.	
		XX XX	
		KW Human; PRO; immune related disorder; systemic lupus erythematosus; rheumatoid arthritis; osteoarthritis; juvenile chronic arthritis; systemic sclerosis; Sjogren's syndrome; vasculitis; sarcoidosis; autoimmune haemolytic anaemia; autoimmune thrombocytopenia; thyroiditis; diabetes mellitus; renal disease; demyelinating disease; central nervous system; peripheral nervous system; demyelinating polyneuropathy; Guillain-Barre syndrome; chronic inflammatory demyelinating polyneuropathy.	
SQ	Sequence 766 AA:	OS Homo sapiens.	
Qy	Query Match 100.0%; Score 3877; DB 8; Length 766;	XX XX	
Qy	Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;	PN WO2004043361-A2.	
Qy	Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	XX XX	
Db	1 NTYRLKLYSLRWSIDHEVLYKQENNLVNAEYGNSSYFLENSTFDEFGHSINDYISPD 60	PD 27-MAY-2004.	
Db	51 NTYRLKLYSLRWSIDHEVLYKQENNLVNAEYGNSSYFLENSTFDEFGHSINDYISPD 110	XX XX	
Qy	61 GQFILLENYVKWQRHSYTASYDIDYDNLNRQLITEERI PNNTOWTVTNSPGHKLAYWN 120	PR 06-NOV-2003; 2003WO-US035268.	
Qy		XX XX	
		PA (GETH) GENENTECH INC.	

XX	Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM; Wood WI, Wu TD;	Qy	601 IAVAPVSRWEYDSVYSTERMGLPPEDNLHYRNSTVMSRAENPKQVBYLLIGTADDN 660
PI	WPI: 2004-420067/39. N-PSSD; ADO19397.	Db	651 IAVAVSRVEYYDSVYSTERMGLPPEDNLHYRNSTVMSRAENPKQVBYLLIGTADDN 710
XX	Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for treating an immune related disorder such as systemic lupus erythematosus or rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or spondyloarthropathy.	Qy	661 VHFOQAQISKALDVGVDFQAMWTDDEGIIASSTAHQHIIYTMASHPTIKQCPLP 716
PT		Db	711 VHFOQAQISKALDVGVDFQAMWTDDEGIIASSTAHQHIIYTMASHPTIKQCPLP 766
PR			RESULT 11
PR			ADO19806
PR			ID ADO19806 standard; protein; 766 AA.
XX		XX	
PS	Claim 7: SEQ ID NO 328; 1731pp; English.	AC	ADO19806;
XX		XX	
CC	The invention relates to human PRO polypeptides and the polynucleotides encoding them. The polypeptides and polynucleotides are useful for treating and diagnosing immune related disorders in mammals. The immune related disorders include systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis, systemic sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune thrombocytopenia, diabetes mellitus, immune-mediated renal disease, demyelinating diseases of the central or peripheral nervous system, demyelinating polyneuropathy, Guillain-Barre syndrome and chronic inflammatory demyelinating polyneuropathy. This sequence represents a human PRO polypeptide of the invention.	DT	12-AUG-2004 (first entry)
CC		XX	Human PRO polypeptide #365.
CC		XX	Human; PRO; immune related disorder; systemic lupus erythematosus; rheumatoid arthritis; osteoarthritis; juvenile chronic arthritis;
CC		XX	KW systemic sclerosis; Sjogren's syndrome; vasculitis; sarcoidosis; autoimmune haemolytic anaemia; autoimmune thrombocytopenia; thyroiditis; diabetes mellitus; renal disease; demyelinating disease; central nervous system; peripheral nervous system; demyelinating polyneuropathy; Guillain-Barre syndrome; chronic inflammatory demyelinating polyneuropathy.
CC		XX	Homo sapiens.
CC		OS	
CC		XX	WO2004043361-A2.
CC		XX	27-MAY-2004.
CC		PD	
CC		XX	06-NOV-2003; 2003WO-US035268.
CC		XX	08-NOV-2002; 2002US-0425235P.
CC		PA	(GETH) GENENTECH INC.
CC		XX	Pong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;
CC		PI	Wood WI, Wu TD;
CC		XX	WPI; 2004-420067/39.
CC		DR	N-PRDB; ADO19805.
CC		XX	Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for treating an immune related disorder such as systemic lupus erythematosus.
CC		PT	PT treating and diagnosing immune related disorders in mammals. The immune related disorders include systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis, systemic sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal disease, demyelinating diseases of the central or peripheral nervous system, demyelinating polyneuropathy, Guillain-Barre syndrome and chronic inflammatory demyelinating polyneuropathy. This sequence represents a human PRO polypeptide of the invention.
CC		XX	Claim 7: SEQ ID NO 730; 1731pp; English.
CC		PS	
CC		XX	The invention relates to human PRO polypeptides and the polynucleotides encoding them. The polypeptides and polynucleotides are useful for treating and diagnosing immune related disorders in mammals. The immune related disorders include systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis, systemic sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal disease, demyelinating diseases of the central or peripheral nervous system, demyelinating polyneuropathy, Guillain-Barre syndrome and chronic inflammatory demyelinating polyneuropathy. This sequence represents a human PRO polypeptide of the invention.
CC		SQ	Sequence 766 AA;
CC		Query Match	100.0%; Score 3877; DB 8; Length 766;
CC		Best Local Similarity	100.0%; Pred. No. 0;
CC		Matches	Mismatches 0; Indels 0; Gaps 0;
CC		Qy	1 NTYRLKYSLRWISDELYKQENNLYFVAEYGNSSVFLENSTDEFGRHSINDYSPD 60
CC		Db	51 NTYRLKYSLRWISDELYKQENNLYFVAEYGNSSVFLENSTDEFGRHSINDYSPD 110
CC		Qy	61 GQFILEEYNTKQWRHSYTASYDIYDLNKRQLITERIPNNTONTWSPGHKLAYWN 120
CC		Db	111 GQFILEEYNTKQWRHSYTASYDIYDLNKRQLITERIPNNTONTWSPGHKLAYWN 170
CC		Qy	121 DIIYKIEPNLPSYRTWTGEDIINGIDWVYEEFVSAYSALWSPNGTFLAYAQFND 180
CC		Db	171 DIIYKIEPNLPSYRTWTGEDIINGIDWVYEEFVSAYSALWSPNGTFLAYAQFND 230
CC		Qy	181 TEVPLIEFSYPSIDESQYPTKTVRYPKAGAVNPTVKFVNTDLSVSYNTATQITAP 240
CC		Db	231 TEVPLIEFSYPSIDESQYPTKTVRYPKAGAVNPTVKFVNTDLSVSYNTATQITAP 290
CC		Qy	241 ASMLIGDHYLCDVTWATQERISLQNLIRQIYNSYMDICDODESSGRNLVARQHIEST 300
CC		Db	291 ASMLIGDHYLCDVTWATQERISLQNLIRQIYNSYMDICDODESSGRNLVARQHIEST 350
CC		Qy	301 TGWGRPRPSEPHFTLDGNSFYK1ISNEGYRHCYQFDKDCTPITKTWEVIGEAL 360
CC		Db	351 TGWGRPRPSEPHFTLDGNSFYK1ISNEGYRHCYQFDKDCTPITKTWEVIGEAL 410
CC		Qy	361 TSDLYYYISNEYKGFMGRNLVYKIQSLDKLRLVLEDNSALDKLQNMQWSKKLDPFLINNETKFWYQML 420
CC		Db	411 TSDLYYYISNEYKGFMGRNLVYKIQSLDKLRLVLEDNSALDKLQNMQWSKKLDPFLINNETKFWYQML 470
CC		Qy	421 RCGSGPGLPLTTLHSSYNDKGLRVEDNSALDKLQNMQWSKKLDPFLINNETKFWYQML 480
CC		Db	471 RCGSGPGLPLTTLHSSYNDKGLRVEDNSALDKLQNMQWSKKLDPFLINNETKFWYQML 530
CC		Qy	481 PPHFDKSKKYPLLDDVAGPCSQADTVFLRNWATYLASTENIVASFDORGSSYQDKI 540
CC		Db	531 PPHFDKSKKYPLLDDVAGPCSQADTVFLRNWATYLASTENIVASFDORGSSYQDKI 590
CC		Qy	541 MHAINRRLGTFEVEDQEARQFSRNGFVDNKRIAIWGSYGGYTTSMVLGSGGVFECG 600
CC		Db	591 MHAINRRLGTFEVEDQEARQFSRNGFVDNKRIAIWGSYGGYTTSMVLGSGGVFECG 650
CC		Query Match	100.0%; Score 3877; DB 8; Length 766;
CC		Best Local Similarity	100.0%; Pred. No. 0;
CC		Matches	Mismatches 0; Indels 0; Gaps 0;

Db 51 NTYRLKLYSLRWSDEHYLKQENNIVFNAEYGNSSVYLENSTDEFGHSINDYSTSPD 110
 QY 61 GQFILEEYNTVKQWRHSHSYTASYDIDYLNRQLITEERIPNNTQWVTSPVGHKLAYWN 120
 Db 111 GQFILEEYNTVKQWRHSHSYTASYDIDYLNRQLITEERIPNNTQWVTSPVGHKLAYWN 170
 QY 121 DIIYKIEPNLPSYRITWTGKEDITYNGTIDWVYBEEFVSAYSALWNSPNTFLAYAQND 180
 Db 171 DIIYKIEPNLPSYRITWTGKEDITYNGTIDWVYBEEFVSAYSALWNSPNTFLAYAQND 230
 QY 181 TEVPLIEYSYPSDESQYPKTVRVPYKAGAVNPYVKKFVNNTDSLSSYNTNATSIQTAP 240
 Db 231 TEVPLIEYSYPSDESQYPKTVRVPYKAGAVNPYVKKFVNNTDSLSSYNTNATSIQTAP 290
 QY 241 ASMLIGDHYLCDVTVATQERISLOWLRQLQNSYMDICDYDESSGRWNLVAQHIENT 300
 Db 291 ASMLIGDHYLCDVTVATQERISLOWLRQLQNSYMDICDYDESSGRWNLVAQHIENT 350
 QY 301 TGWYGRFRPSEPHFTLDGNSFYKLTISNEEGYRHICFYQIDKUDCTFTKGTVWEIGRAL 360
 Db 351 TGWYGRFRPSEPHFTLDGNSFYKLTISNEEBYRHICFYQIDKUDCTFTKGTVWEIGRAL 410
 QY 361 TSDLYYYISNEYKMPGGRNLKYLQSLDYTKVCLSCLENPERQOYYSVSFSKEAKYQL 420
 Db 411 TSDLYYYISNEYKMPGGRNLKYLQSLDYTKVCLSCLENPERQOYYSVSFSKEAKYQL 470
 QY 421 RCGFGLPLTYLHSSVNDKGRLRVEDNSALDKMQLQNVOMPSSKLLDFILINETKFWYQNL 480
 Db 471 RCGFGLPLTYLHSSVNDKGRLRVEDNSALDKMQLQNVQMPSSKLLDFILINETKFWYQNL 530
 QY 481 PPHFDKSKYCKYPLLIDVYAGFCOSOKADTVFRLNWATYLAESTENITVASFDRGSGSYQDKI 540
 Db 531 PPHFDKSKYCKYPLLIDVYAGCSOKADTVFRLNWATYLAESTENITVASFDRGSGSYQDKI 590
 QY 541 MHAINRRLGTPEVBDQIEARQESKMGFDNKRIAIWGRSISGGYVTSMVLGSSGVPRCG 600
 Db 591 MHAINRRLGTPEVBDQIEARQESKMGFDNKRIAIWGRSISGGYVTSMVLGSSGVPRCG 650
 QY 601 IAVAPVSRWEYDSYTERYMGLLPTPEDNLDHYRNSTMRAENFKQEYLLHGADDN 660
 Db 651 IAVAPVSRWEYDSYTERYMGLLPTPEDNLDHYRNSTMRAENFKQEYLLHGADDN 710
 QY 661 VHFOQSAAIQSKALVDVGDFQAMMYTDEHDGIASTAHQIYTHMSFIKQCFSLP 716
 Db 711 VHFOQSAAIQSKALVDVGDFQAMMYTDEHDGIASTAHQIYTHMSFIKQCFSLP 766

RESULT 12

AD071612 standard; protein; 766 AA.

XX PI Dang NH, Morimoto C;
 XX DR WPI: 2004-420511/39.
 XX DR N-PSDB; AD071611, AD071613.
 Use of a CD26 composition, and a chemotherapeutic and/or a radiotherapeutic agent for e.g. inhibiting the cell growth, inducing cell cycle arrest, killing a cancer cell, treating cancer, or inducing tumor regression or tumor necrosis.
 XX Claim 23; Page 151-153; 182pp; English.
 XX The specification describes a CD26 composition which, in conjunction with a chemotherapeutic or radiotherapeutic agent, is used for the treatment of cancers. Expression of CD26 enhances the sensitivity of the cancer cell to the chemotherapeutic or radiotherapeutic agent. CD26 is a dipeptidyl peptidase IV (DPPIV). The hemotherapeutic agent is topoisomerase II inhibitor. The CD26 composition of the invention is useful for inhibiting the growth of a cell, inducing cell cycle arrest in a cell, killing a cancer cell, potentiating the effect of a chemotherapeutic agent and/or a radiotherapeutic agent on a tumour cell, inducing or enhancing apoptosis of a cancer cell, treating cancer, or inducing tumor regression or tumour necrosis. The CD26 composition is further useful for increasing topoisomerase II expression in a cell, for activating an antigen-presenting cell, or for potentiating immune responses of an animal. The present sequence represents a CD26 protein, and is encoded by vectors which are used to produce compositions of the invention.
 XX Sequence 766 AA;
 SQ Query Match 100.0%; Score 3877; DB 8; Length 766;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 NTYRLKLYSLRWSDEHYLKQENNIVFNAEYGNSSVYLENSTDEFGHSINDYSTSPD 60
 Db 51 NTYRLKLYSLRWSDEHYLKQENNIVFNAEYGNSSVYLENSTDEFGHSINDYSTSPD 110
 Qy 61 GQFILEEYNTVKQWRHSHSYTASYDIDYLNRQLITEERIPNNTQWVTSPVGHKLAYWN 120
 Db 111 GQFILEEYNTVKQWRHSHSYTASYDIDYLNRQLITEERIPNNTQWVTSPVGHKLAYWN 170
 Qy 121 DIVYKIEPNLPSYRITWTGKEDITYNGTIDWVYBEEFVSAYSALWNSPNTFLAYAQND 180
 Db 171 DIVYKIEPNLPSYRITWTGKEDITYNGTIDWVYBEEFVSAYSALWNSPNTFLAYAQND 230
 Qy 181 TEVPLIEYSYPSDESQYPKTVRVPYKAGAVNPYVKKFVNNTDSLSSVNTNATSIQTAP 240
 Db 231 TEVPLIEYSYPSDESQYPKTVRVPYKAGAVNPYVKKFVNNTDSLSSVNTNATSIQTAP 290
 Qy 241 ASMLIGDHYLCDVTVATQERISLOWLRQLITEERIPNNTQWVTSPVGHKLAYWN 300
 Db 291 ASMLIGDHYLCDVTVATQERISLOWLRQLITEERIPNNTQWVTSPVGHKLAYWN 350
 Qy 301 TGWYGRFRPSEPHFTLDGNSFYKLTISNEEGYRHICFYQIDKUDCTFTKGTVWEIGRAL 360
 Db 351 TGWYGRFRPSEPHFTLDGNSFYKLTISNEEGYRHICFYQIDKUDCTFTKGTVWEIGRAL 410
 Qy 361 TSDLYYYISNEYKMPGGRNLKYLQSLDYTKVCLSCLENPERQOYYSVSFSKEAKYQL 420
 Db 411 TSDLYYYISNEYKMPGGRNLKYLQSLDYTKVCLSCLENPERQOYYSVSFSKEAKYQL 470
 Qy 421 RCSPGGLPLYTHSVDKGRLRVEDNSALDKMQLQNVOMPSSKLLDFILINETKFWYQNL 480
 Db 471 RCSPGGLPLYTHSVDKGRLRVEDNSALDKMQLQNVOMPSSKLLDFILINETKFWYQNL 530
 Qy 481 PPHEFDKSKYCKYPLLIDVYAGPCSKADTVFRLNWATYLASTENITVASFDRGSGSYQDGKI 540
 S.1 PPHEFDKSKYCKYPLLIDVYAGPCSKADTVFRLNWATYLASTENITVASFDRGSGSYQDGKI 590
 PA (TEXA) UNIV TEXAS SYSTEM.

QY	591	MHAIRRGLTPEYEDQLEBAAQFSKMGIVDNKRIAIRGSWSYGGVTSNVLGSSSGVPRCG	650	Db	1	NTVPLKLYSLRWISDHEVLYKOENNILVNAEYGNSSVLENSTEDEFGHISINDYSPD
Qy	601	IAVAPVSEWEYDSVYTRYMLGPTPEINLDHYTRNSTMSRAINFQKVYLLHGADDN	660	Db	51	NTVPLKLYSLRWISDHEVLYKOENNILVNAEYGNSSVLENSTEDEFGHISINDYSPD
Qy	651	IAVAPVSEWEYDSVYTRYMLGPTPEINLDHYTRNSTMSRAINFQKVYLLHGADDN	710	Db	61	GQFILLEMYNVVKQRHHSYTAQSYDYLNRQLITEERIPNNTQTYTWSPYGHKLAYWN
Qy	661	VHFQQSAQISKAVALDVYGDQFOAMWYTDDEDHGIASSTAQOHITHMSHPIKOCFLP	716	Db	111	GQFILLEMYNVVKQRHHSYTAQSYDYLNRQLITEERIPNNTQTYTWSPYGHKLAYWN
db	711	VHFQQSAQISKAVALDVYGDQFOAMWYTDDEDHGIASSTAQOHITHMSHPIKOCFLP	766	Db	121	DIYKIEPNLPSYLTWTGKEDIYNGTIDWVYEEPVFSAWJWSPNTFLAYAQFND
Db				Db	171	DIYKIEPNLPSYLTWTGKEDIYNGTIDWVYEEPVFSAWJWSPNTFLAYAQFND
Qy				Qy	181	TEVPLIEYSFYSDSLSQYPKTVRYPKAGAVNPYKTFVYNTDSLSSVTNATSIOTAP
XX				Db	231	TEVPLIEYSFYSDSLSQYPKTVRYPKAGAVNPYKTFVYNTDSLSSVTNATSIOTAP
AC				Qy	241	ASMLIGDHVLCDTWATQBRISLWLRRLQNYSMIDICDYOESCRWNLVQARQHLEMST
XX				Db	291	ASMLIGDHVLCDTWATQBRISLWLRRLQNYSMIDICDYOESCRWNLVQARQHLEMST
DB				Qy	301	TGWGRFRSEPHFTLDGNSFYKIIISNEEGYRHICFYQFQIDKCDTPTKGTWVIGEAL
DB				Db	351	TGWGRFRSEPHFTLDGNSFYKIIISNEEGYRHICFYQFQIDKCDTPTKGTWVIGEAL
KW				Qy	361	TSDLYYISNEYKGMPGGERNLKYKIQLDSDTYKTVTCLSCELNPERCQYSYSFSKEAR
KW				Db	411	TSDLYYISNEYKGMPGGERNLKYKIQLDSDTYKTVTCLSCELNPERCQYSYSFSKEAR
XX				Qy	421	RCSPGGLPLTYLTHSVNDKGLRVLIEDNSALDKMILQVNQMFSPSKKDFIILNETKPYQML
XX				Db	471	RCSPGGLPLTYLTHSVNDKGLRVLIEDNSALDKMILQVNQMFSPSKKDFIILNETKPYQML
XX				Qy	481	PPHPFDKSKKXPLPLDVAAGPCSOAKADTYVERLNWATYLASTENIVVASFDGRGSGYQDKI
XX				Db	531	PPHPFDKSKKXPLPLDVAAGPCSQADTYVERLNWATYLASTENIVVASFDGRGSGYQDKI
PD				Qy	541	MHAIRRGLTPEYEDQIEAROFKMGFVDNKRFAIWGMNSYGGVUTSMVILGSGGVFKCG
XX				Db	591	MHAIRRGLTPEYEDQIEAROFKMGFVDNKRFAIWGMNSYGGVUTSMVILGSGGVFKCG
PF				Qy	601	IAYVPVSRVEYYDVSYTETYMGGLPTPEDNLHDYHNSTVNSRAENFRQVEYLILLGTA
XX				Db	651	IAYVPVSRVEYYDVSYTETYMGGLPTPEDNLHDYHNSTVNSRAENFRQVEYLILLGTA
XX				Qy	661	VHFQQAQSIAQSKALVDVGIVDFQAMWYTDDEDHGIASSTAHOHYTMHSHP1KQCFSLP
PR				Db	711	VHFQQAQSIAQSKALVDVGIVDFQAMWYTDDEDHGIASSTAHOHYTMHSHP1KQCFSLP
XX				XX		RESULT 14
XX				XX		ABM80355
XX				XX		ID ABM80355 standard; protein; 766 AA.
XX				XX		AC ABM80355;
XX				DT 18-NOV-2004 (first entry)		
XX						Tumour-associated antigenic target (TAT) polypeptide PROB0881, SEQ: 895.
XX						The specification describes a CD26 composition which, in conjunction with a chemotherapeutic or radiotherapeutic agent, is used for the treatment and prevention of cancers. Expression of CD26 enhances the sensitivity of the cancer cell to the chemotherapeutic or radiotherapeutic agent; CD26 is a dipeptidyl peptidase IV (DPP4). The chemotherapeutic agent is a topoisomerase II inhibitor. The CD26 composition of the invention is useful for inhibiting the growth of a cell, inducing cell cycle arrest in a cell, killing a cancer cell, potentiating the effect of a chemotherapeutic agent and/or a radiotherapeutic agent on a tumour cell, inducing or enhancing apoptosis of a cancer cell, treating cancer, or inducing tumour regression or tumour necrosis. The CD26 composition is further useful for increasing topoisomerase II expression in a cell, for activating an antigen-presenting cell, or for potentiating immune responses of an animal. The present sequence represents a CD26 protein, and is encoded by vectors which are used to produce compositions of the invention.
XX				XX		XX Tumour-associated antigenic target; TAT; human; overexpression; cancer; tumour; diagnosis; cell proliferative disorder; breast cancer; colorectal cancer; lung cancer; ovarian cancer; liver cancer; central nervous system cancer; bladder cancer; pancreatic cancer; cervical cancer; melanoma; leukemia; hybridisation probe; chromosome identification; cytostatic; gene mapping; gene therapy; cytotoxicity.
XX				XX		XX Homo sapiens
XX				XX		Sequence 766 AA;

comprising a container, a label on the container and a composition of
 CC matter of (7); (9) a method of treating an immune related disease in a
 CC mammal; (10) a method for determining the presence of a PRO polypeptide
 CC in a sample suspected of having the polypeptide; (11) a method of
 CC diagnosing an immune related disease or an inflammatory
 CC response in a mammal; (12) a method of identifying a compound that
 CC inhibits or mimics the activity of or expression of gene encoding
 CC ; and (13) a method of stimulating the immune response in a mammal. The
 CC PRO sequences have antiallergic, antianæmic, antiarbititic,
 CC antiasthmatic, antidiabetic, antiinflammatory, antipsoriatic,
 CC antirheumatic, antithyroid, CNS, dermatological, gastrointestinal,
 CC haemostatic, hepatotropic, immunomodulant, immunosuppressive, muscular,
 CC nephrotropic, neuroprotective, osteopathic, respiratory, vasotrophic and
 CC virucide activities, and can be used in gene therapy. The nucleic acid
 CC (I) and the encoded polypeptides, compositions, kits and methods are
 CC useful in diagnosing and treating an immune related disease and in
 CC stimulating an immune response. The present sequence represents a human
 CC PRO protein from the present invention.

Sequence 766 AA;

	Query Match	Score	Length	DB	Match	Best Local Similarity	Pred.	No.	Mismatches	Indels	Gaps	O;
Qy	1 NTYRLKLYSURWISDHLYLKQENNILVFNABGNSSVFULENSTDEFHSINDYSISPD	60	716	Matches	716; Conservative	100 %	0;	0	0;	0	0	
Db	51 NTYRLKLYSLRWISEHLYLKQENNILVFNABGNSSVFULENSTDEFHSINDYSISPD	60										
Qy	61 QQFILEEYNTVKQRHSYTASYDIYDLNRQLITERIPNNTQWNTWPVGHKLAYWN	120										
Db	111 QQFILEEYNTVKQRHSYTASYDIYDLNRQLITERIPNNTQWNTWPVGHKLAYWN	120										
Qy	121 DIVYKLEPNLPSYRTWTGKEDIYINGITDWYVSEEVSYASALMWSPIGTFLAYAQEND	180										
Db	171 DIVYKLEPNLPSYRTWTGKEDIYINGITDWYVSEEVSYASALMWSPIGTFLAYAQEND	180										
Qy	181 TEVPLIBSFYSDESLOYPKIVRVPYPRAGAVNPVFKFVVNTDSLSVTNATSIQTAP	240										
Db	231 TEVPLIEFSYSDESLQYPKIVRVPYPRAGAVNPVFKFVVNTDSLSVTNATSIQTAP	240										
Qy	241 ASMLIGDHYLCDVTWATQERISLQWLRIRONYSYMIDCYDESSGRWNLVAROHIEST	300										
Db	291 ASMLIGDHYLCDVTWATQERISLQWLRIRONYSYMIDCYDESSGRWNLVAROHIEST	300										
Qy	301 TGWGRFRPSEPHFTLDGNSFYKLTISNEECYRICYFQIDKKDCTFITKGTWEGIENL	360										
Db	351 TGWGRFRPSEPHFTLDGNSFYKLTISNEECYRICYFQIDKKDCTFITKGTWEGIENL	360										
Qy	361 TSDILYYISNEYKGMPGNLYKIQLSDTKVLCLSCLENPERQYYSFSREAKYVQL	420										
Db	411 TSDILYYISNEYKGMPGNLYKQLSDTCKVLCLSCLENPERQYYSFSREAKYVQL	420										
Qy	421 RCGFGLPLTYLHSSVNDKGRLVIEDNSALDKLQNVMQPSKCLDFILNETKFWYQML	480										
Db	471 RCGFGLPLTYLHSSVNDKGRLVIEDNSALDKLQNVMQPSKCLDFILNETKFWYQML	480										
Qy	481 PPHEFIKSCKYPLLIIDVYAGPCSKQADTVPLRNWATYLASENITIVASPGRGSSYQGKX	540										
Db	531 PPHEFIKSCKYPLLIIDVYAGPCSKQADTVPLRNWATYLASENITIVASPGRGSSYQGKX	540										
Qy	541 MHAIRRLGFEVEDQEARQFSKMGFDNKRAIwgSYGGVTSVTLGSGGVFTCG	600										
Db	591 MHAIRRLGFEVEDQEARQFSKMGFDNKRAIwgSYGGVTSVTLGSGGVFTCG	600										
Qy	601 IAVAPVSRNEYDSYTERMGLPLTPEDNLHDYNSTMRAENFKOYEYLHGTADDN	660										
Db	651 IAVAPVSRNEYDSYTERMGLPLTPEDNLHDYNSTMRAENFKOYEYLHGTADDN	660										
Qy	661 VHFQOSAQSIAKLVDFQAMXYTDEDHGIASTAIQHITYTMSHFTKQCSLP	716										
Db	711 VHFQOSAQSIAKLVDFQAMXYTDEDHGIASTAIQHITYTMSHFTKQCSLP	716										

Search completed: January 30, 2006, 14:22:35
 Job time : 138 secs

THIS PAGE BLANK (USP1c)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

M protein - protein search, using sw model
run on: January 30, 2006, 14:14:31 : Search time 41 Seconds
(without alignments)
1680.273 Million cell updates/sec

title: US-10-659-055-1_COPY_51_766
effect score: 38.77
sequence: 1 NTYRLKLYSLRWSIDHEVLY.....AHQIYTHMSHF1KQCFSIP 716
scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

searched: 283416 seqs, 96216763 residues
total number of hits satisfying chosen parameters: 283416

post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

database : PIR:80:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	DB ID	Description
1	3871	99.8	766	CDH726 dipeptidyl-peptida
2	3353.5	86.5	792	1 A39314 dipeptidyl-peptida
3	3336	86.0	760	1 S23152 dipeptidyl-peptida
4	1930.5	49.8	759	2 I38893 fibroblast activat
5	1117	28.8	803	2 I68800 dipeptidyl aminope
6	1117	28.8	865	2 I54331 dipeptidyl aminope
7	1094	28.2	803	2 A41193 dipeptidyl aminope
8	940	24.2	793	2 T410103 dipeptidyl aminope
9	908	23.4	818	1 S4007 dipeptidyl aminope
10	791.5	20.4	711	2 S66261 X-Pro dipeptidyl-p
11	745	19.2	829	2 T19314 hypothetical prote
12	726	18.7	931	2 A49137 dipeptidyl aminope
13	720	18.6	779	2 T25173 hypothetical prote
14	716	18.5	799	2 T25174 hypothetical prote
15	628	16.2	738	2 JC87156 dipeptidyl peptide
16	594.5	15.3	741	2 S66261 X-Pro dipeptidyl-p
17	590.5	15.2	743	2 T19314 probable dipeptidyl
18	445.5	11.5	931	2 T37700 hypothetical prote
19	332	8.6	795	2 F82858 dipeptidyl-peptida
20	272	7.0	657	2 E70255 probable acylamino
21	244.5	6.3	738	2 JC87156 acylaminoacyl-pept
22	222.5	5.7	622	2 F71174 hypothetical prote
23	221	5.7	591	2 H72474 probable acylamino
24	219	5.6	642	2 C71137 hypothetical prote
25	206.5	5.3	632	2 E715057 peptidease PAB14196
26	205	5.3	631	2 H75007 probable acylamino
27	205	5.3	683	2 B87495 prolyl oligopeptid
28	191	4.9	709	2 B87280 alanyl dipeptidyl
29	188	4.8	659	2 F72568 probable acylamino

RESULT 1						
CDHU26 dipeptidyl-peptidase IV (EC 3.4.14.5) - human						
N;Alternate names: cell surface glycoprotein CD26; thymocyte-activating molecule						
C;Species: Homo sapiens (man)						
C;Date: 31-Dec-1993 #sequence revision 23-Aug-1996 #text change 09-Jul-2004						
C;Accession: S24313; B42408; #42408; B61136; S59510; I56154; S59857; S15520						
R;Mizumi, Y.; Hayashi, Y.; Arakawa, F.; Ichihara, Y.						
Biochim. Biophys. Acta 1131: 333-336, 1992						
A;Title: Molecular cloning and sequence analysis of human dipeptidyl peptidase IV mRNA levels during cell differentiation.						
A;Reference number: S24313; MUID:92329951; PMID:1352704						
A;Accession: S24313						
A;Molecule type: mRNA						
A;Residues: 1-6, 'I'-8-766 <MS>						
A;Cross-references: UNIPROT:U007487; UNIPARC:UPI000016AE80; EMBL:X60708; NID:915						
R;Darmoul, D.; Lacasa, M.; Baricault, L.; Marguet, D.; Sappi, C.; Trotot, P.; B						
J. Biol. Chem. 267: 4824-4833, 1992						
A;Title: Dipeptidyl peptidase IV (CD 26) gene expression in enterocyte-like col						
A;Reference number: A42408; MUID:92165547; PMID:1347043						
A;Accession: B42408						
A;Molecule type: protein						
A;Residues: 1-5, 'R'-7-436, 'S'-438-556, 'I', 558-662, 'E', 664-766 <DAR1>						
A;Cross-references: UNIPARC:UPI0000052ACB; GB:MB0536; PIDN:AAA5230						
A;Experimental source: intersine						
A;Note: this sequence corresponds with the author's translation						
A;Accession: A42408						
A;Molecule type: mRNA						
A;Residues: 1-5, 'R'-7-436, 'S'-438-556, 'I', 558-662, 'E', 664-766 <DAR1>						
A;Cross-references: UNIPARC:UPI0000172ACB; GB:MB0536; PIDN:9181569						
A;Note: sequence extracted from NCBI backbone (NCBIN:83986; NCNP:82988); this						
R;Gorvel, J.P.; Ferrero, A.; Chambraud, L.; Rigal, A.; Bonicel, J.; Maroux, S.						
Gastroenterology 101: 618-625, 1991						
A;Title: Expression of sucrase-isomaltase and dipeptidylpeptidase IV in human s						
A;Reference number: A61136; MUID:91317403; PMID:1677636						
A;Accession: B61136						
A;Molecule type: protein						
A;Residues: 1-15, 'X'-17-22 <GOR>						
A;Cross-references: UNIPARC:UPI000016BA6; GB:S79876; NID:91195574; PIDN:AA8356						
R;Boehm, S.K.; Gum Jr., J.R.; Erickson, R.H.; Hicks, J.W.; Kim, Y.S.						
Biochem. J. 311: 835-843, 1995						
A;Title: Human dipeptidyl peptidase IV gene promoter: tissue-specific regulation						
A;Reference number: I56154; MUID:96067599; PMID:7487939						
A;Accession: I56154						
A;Molecule type: DNA						
A;Residues: 1-31 <BOE>						
A;Cross-references: UNIPARC:UPI0000052ACB						
R;Tanaka, T.; Camerini, D.; Seed, B.; Torimoto, Y.; Dang, N.H.; Kameoka, J.; Da						
J. Immunol. 149: 481-486, 1992						
A;Title: Cloning and functional expression of the T cell activation antigen CD2						
A;Reference number: I56154; MUID:92323476; PMID:1352530						

A;Status: translated from GB/EMBL/DDJB		Qy	661 VHFQQSAQISKALVLDVGVDQAMWYTDDEHGIASSTAHQIHYTHMSHFIKQCFSLP 716
A;Molecule type: mRNA		Db	711 VHFQQSAQISKALVLDVGVDQAMWYTDDEHGIASSTAHQIHYTHMSHFIKQCFSLP 766
A;Residues: 1-436; 'S', 438-766 <TAN>			
A;Cross-references: UNIPARC:UPI000004F7BF; GB:MT4777; NID:9180082; PIDN:AAA51943_1; PMID:R; Abbott, C.A.; Baker, E.; Sutherland, G.R.; McCaughan, G.W.			
R; Abbott, C.A.; Baker, E.; Sutherland, G.R.; McCaughan, G.W.			
A;Title: Genomic organization, exact localization, and tissue expression of the human CD1A gene			
A;Reference number: S53857; MUID:05012454; PMID:792537			
A;Accession: S53857			
A;Molecule type: DNA			
A;Residues: 1-436; 'S', 438-766 <ABB>			
A;Cross-references: UNIPARC:UPI000004F7BF; EMBL:U13734			
C;Genetics:			
A;Gene: GDB:DPP4			
A;Cross-references: GDB:125239; OMIM:102720			
A;Map position: 2G14.3-2q24.3			
A;Introns: 2/3; 32/1; 65/1; 95/3; 122/3; 140/2; 164/3; 205/1; 258/3; 296/2; 341/3; 356/3			
C;Keywords: dipeptidyl-peptidase IV, glycoprotein, homodimer, proteinase, transmembrane domain, intracellular #status predicted <INT>			
P;7/28/Domain: transmembrane #status predicted <TRN>			
P;29-766/Domain: extracellular #status predicted <EXT>			
P;85, 92, 150, 219, 229, 281, 321, 520, 685/Binding site: carbohydrate (Asn) (covalent) #status predicted			
P;630, 708, 740/Active site: Ser, Asp, His #status predicted			
Query Match	99.8%	Score 3871; DB 1; Length 766;	
Best Local Similarity	99.9%	Pred. No. 5.9e-260;	
Matches 715; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy	1 NTYDLKLYSRWISDHEVLYKOENNLIVNAEYNSSYLENSTDFEFGHSINDYSISPD 60	Db	51 NTYDLKLYSRWISDHEVLYKOENNLIVNAEYNSSYLENSTDFEFGHSINDYSISPD 110
Qy	61 QGFILEEYNTVKQRHSTASYDYLNLQRQLTTEREPNTQWTVSPVGHKLAYWN 120	Db	111 QGFILEEYNTVKQRHSTASYDYLNLQRQLTTEREPNTQWTVSPVGHKLAYWN 170
Qy	121 DYYKKIEPNLPSYRITWTGKEDIYNGTIDWYEEEVFSAYSALWWSNGTFLAYAQND 180	Db	171 DYYKIEPNLPSYRITWTGKEDIYNGTIDWYEEEVFSAYSALWWSNGTFLAYAQND 230
Qy	181 TEVPLIEYSFYSDLSQYPKTVRVPYPKAGAVNPTVKFFVNTDSLSSVTNATSIQTAP 240	Db	231 TEVPLIEYSFYSDLSQYPKTVRVPYPKAGAVNPTVKFFVNTDSLSSVTNATSIQTAP 290
Qy	241 ASMLJGDHYLCDTVATQPRISLQLRQLQNSYMDIDCYDESSGRNCLVARQHIESM 300	Db	291 ASMLJGDHYLCDTVATQPRISLQLRQLQNSYMDICDYDESSGRNCLVARQHIESM 350
Qy	301 TGKYGGRFRPSEPHFTLGNSFYKLTISNEGGYRHICYFQDQKDDCTFTKGTVEWIGAL 360	Db	351 TGWGRFRPSEPHFTLGNSFYKLTISNEGGYRHICYFQDQKCTFTKGTVEWIGAL 410
Qy	411 TSDLYYYISNEYKGMPCGRNLQYKIQSLQSYTKVTCSCLENPERCQYTSVSFSKEAKYQ 420	Db	411 TSDLYYYISNEYKGMPCGRNLQYKIQSLQSYTKVTCSCLENPERCQYTSVSFSKEAKYQ 470
Qy	421 RCGSPGLPLIYTLLHSVNQDKLRLVEDNSALDKLQNTVOPSKLDIFILNEKYKFWQNL 480	Db	471 RCGSPGLPLIYTLLHSVNQDKLRLVEDNSALDKLQNTVOPSKLDIFILNEKYKFWQNL 530
Qy	481 PPHFDKSXKCYPLLIQDVAQPCSKQADTYFRLNWATYLASTENITIVASFDFGRSGYQDKI 540	Db	531 PPHFDKSXKCYPLLIQDVAQPCSKQADTYFRLNWATYLASTENITIVASFDFGRSGYQDKI 590
Qy	541 MHA1NRRLGTFEVDQIEAROFRSKMGFVNKRIAIWGSYGGVTSWVLSGSSGVFKCG 600	Db	591 MHA1NRRLGTFEVDQIEAROFRSKMGFVNKRIAIWGSYGGVTSWVLSGSSGVFKCG 650
Qy	601 IAVAPVSRMEYYDSVYTERYMLGFLPTPEVDLHYRNSTMRAENFKQTYBLLIHGTADDN 660	Db	651 IAVAPVSRMEYYDSVYTERYMLGFLPTPEVDLHYRNSTMRAENFKQTYBLLIHGTADDN 710

P;1-28/Domain: signal sequence #link MATS #status experimental <SIG>
 P;1-6/Domain: intracellular #status predicted <INT>
 P;7-28/Domain: transmembrane #status predicted <TMN>
 P;29-34/Domain: extracellular #status predicted <EXT>
 P;35-792/Domain: protein #link MATS #status experimental <PRO>
 P;35-792/Product: dipeptidyl-peptidase, soluble form #status experimental
 P;631/Active site: carbohydrate (Asn) (covalent) #status predicted
 P;709,741/Active site: Asp, His #status predicted

Query Match 86.5%; Score 3153.5; DB 1; Length 792;
 Best Local Similarity 84.5%; Pred. No. 4.e-224;
 Matches 607; Conservative 53; Mismatches 55; Indels 3; Gaps 1;

Qy 1 NTYRLKLYSURWISDHELYKQENNLVFNAYGNSSVPLDENSTEDEFHSINDYSISPD 60
 Db 49 NTERTKSYKSRWVSESELYKQENNLVFNAYGNSSVPLDENSTEDEFHSINDYSISPD 108

Qy 61 GQFILEEYNTVKWHRHSTASYDYLNLQRQLTEERIPNTQWTVTWSPVGHKLAYVN 120
 Db 109 RLFLVLLYNTVKWHRHSTASYDYLNLQRQLTEEKIPNTQWTVTWSPVGHKLAYVN 168

Qy 121 DIVYKIEPNLPSRITWTGKEDIYNGTIDWYEEVEFVAYSALWWSPNGTFLAYAQEND 180
 Db 169 DIVYKIEPHLPSRHTSTGKENVTFNGTDWYEEFVAYSALWWSPNGTFLAYAQEND 228

Qy 181 TEVPLIEYSPYSDESLOQPKTVRVPYPAGAVNPVTKFFVWNNTDSLSVTNTASIQITAP 240
 Db 229 TGVPLIEYSPYSDESLOQPKTVWVPYPAGAVNPVTKFFVWNNTDSLSVTNTAP 288

Qy 241 ASMLJGDHYLCDFDTWATQERISLWRTRIONYSMDICDYEDESSGRMNCVLAZHENT 300
 Db 289 ASVITGDHYLCDFDAWSEDISLWRTRIONYSMDICDYDCKTNLWVNPOTQHIESA 348

Qy 301 TGWYGRPRSEPHFTLDGNSFPYKLTISNEGYRHCYFQLDKK--DCPRTIKGTWEVIGI 357
 Db 349 TGWYGRPRPAEPHTSDGSSFPYKLVSDKDCYKHCQFDKRPKEQDCFTIKGCAEWVSI 408

Qy 358 EALTSDLYYISNEYKGMPEGRNLYKIQSLDYTKTCSCELNPERCQYYSVSFSKEARY 417
 Db 409 EALTSDLYYISNEYKEMPGGRNLYKIQLTDHMKCLSDLNPERCQYYSVSLSKEAKY 468

Qy 418 YQRCGSGPGLPLYTIAHSYNDKGKLRVLEDNSALDKMLQVQMPKKEKQFILNETKFYQ 477
 Db 469 YQLGGRGPGLPLYTIAHSRSTDQEKEVRVLEDNSALDKMLQVQMPKKEKQFILNETRFYQ 528

Qy 478 MILPPHPDKSKKKYKPYLDDIVYAGPCSQKADTYVFLNWATYLASTPENIYASFDGRGSYQ 537
 Db 529 MILPPHPDKSKKKYKPYLDDIVYAGPCSQKADAFAFNWATYLASTPENIYASFDGRGSYQ 588

Qy 538 DKIMHAINRRLGTPEVDQLEAAQFSKNGFVDNMKRALIWGWSYGGTTSMVJGSGSVF 597
 Db 589 DKIMHAINRRLGTLEVEDQLEAAQFQLNGFVDSKQVAIWGWSYGGTTSMVJGSGGVF 648

Qy 598 KCGIAAVAPYRSRWEYYDSVTERYMLPTEBDNLHYRNTSTMRAENPKQVEYLIRSTA 657
 Db 649 KCGIAAVAPYRSRWEYYDSVTERYMLPTEBDNLHYRNTSTMRAENPKQVEYLIRSTA 708

Qy 658 DDNYTHFOQSAQISKVALDVGDFQAMWTTDEDHGIASTAHQHITYTMSPHIQCPSL 715
 Db 709 DDNYTHFOQSAQISKVALDVGDFQAMWTTDEDHGIASTAHQHITYSHLHQCPSL 766

Db 750 DDNYTHFOQSAQISKVALDVGDFQAMWTTDEDHGIASTAHQHITYSHLHQCPSL 766

Db 805 TSDYLYTISNEYKGMPGGRNLYKIQSLDYTKTCLSCPBLNPQCQYSYSFSKEAKYQVL 420

Db 823352 dipeptidyl-peptidase IV (EC 3.4.14.5) alpha chain - mouse
 N;Alternate names: CD26 alpha subunit; THM alpha subunit
 C;Species: Mus musculus (house mouse)
 C;Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 10-Sep-1999
 C;Accession: S23732; A46465; A563030
 R;Marguet, D.; Bernard, A.M.; Vivier, I.; Darmoul, D.; Naquet, P.; Pierres, M.
 J. Biol. Chem. 267, 220-228, 1992
 A;Title: cDNA cloning for mouse thymocyte-activating molecule. A multifunctional ecto-di-

A;Reference number: S23752; MUID:92129288; PMID:1370813
 A;Accession: S23752
 A;Status: Preliminary
 A;Molecule type: mRNA
 A;Residues: 1-760 <MAR>
 A;Cross-references: UNIPARC:UPI0000172A2D; EMBL:X58384
 R;Vivier, I.; Marguet, D.; Naquet, P.; Bonicel, J.; Black, D.; Li, C.X.; Bernard, A.M.;
 J. Immunol. 147, 447-454, 1991
 A;Title: Evidence that thymocyte-activating molecule is mouse CD26 (dipeptidyl peptidase
 A;Reference number: A46465; MUID:91302787; PMID:1712807
 A;Accession: A46465
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 1-20 <VIV>
 A;Cross-references: UNIPARC:UPI0000172A2E
 A;Experimental source: M14.7 thymoma cells, Swiss nu/nu
 A;Note: sequence extracted from NCBI backbone (NCBIP:42236)
 R;Bernard, A.M.; Mattei, M.G.; Pierres, M.; Marguet, D.
 Biochemistry 33, 15204-15214, 1994
 A;Title: Structure of the mouse dipeptidyl peptidase IV (CD26) gene
 A;Reference number: A56030; MUID:9509780; PMID:7999781
 A;Accession: A56030
 A;Gene: CD26
 A;Superfamily: dipeptidyl-peptidase IV
 C;Keywords: dipeptidyl-peptidase hydrolase; glycoprotein; transmembrane protein
 F;213,223,315,514,679 Binding site: carbohydrate (Asn) (covalent) #status predicted
 F;624,702,734/Active site: Ser, Asp, His #status predicted
 C;Genetics:
 Qy 1 NTYRLKLYSLRWSIDHLYKQENNLVFNAYGNSSVPLDENSTEDEFHSINDYSISPD 60
 Db 49 STPRVKSYSLWVYSDFEELYKQENNLVFNAYGNSSVPLDENSTEDEFHSINDYSISPD 60
 Db 606; Conservatve 55; Mismatches 50; Indels 4; Gaps 1;
 Query Match 86.0%; Score 3336; DB 1; Length 760;
 Best Local Similarity 84.8%; Pred. No. 6.4e-223;
 Matches 606; Conservatve 55; Mismatches 50; Indels 4; Gaps 1;
 Qy 61 GQPLLBINYVKWHRHSTASYDYLINKRQLITEERIPNTNTWVTPVGHKLAYVN 120
 Db 105 RLFLVLLYNTVKWHRHSTASYDYLINKRQLITEERIPNTNTWVTPVGHKLAYVN 164
 Db 121 DITYKIEBNLPSRITWKGEDIYNGTIDWYEEVEFVAYSALWWSPNGTFLAYAQEND 180
 Db 165 DITYKVBPHLPSRITSGEENYIYNGTIDWYEEVEFVAYSALWWSPNTFLAYAQEND 224
 Db 181 TEVPLIEYSFYSDSDELQPKTVRVPYKAGAVNPVTKFFVWNNTDSLSVTNTASIQITAP 240
 Db 225 TGWYGRPRSEPHFTLDGNSFPYKLTISNEYKGMPEGRNLYKIQSLDYTKTCLSCPBLNPQCQYSYSFSKEAKYQVL 284
 Db 241 ASMLJGDHYLCDFDTWATQERISLWRTRIONYSMDICDYEDESSGRMNCVLAZHENT 300
 Db 285 ASVGDHYLCDFVWATERISLWRTRIONYSMDICDYEDESSGRMNCVLAZHENT 344
 Db 301 TGWYGRPRSEPHFTLDGNSFPYKLTISNEYKEMPGGRNLYKIQSLDYTKTCLSCPBLNPQCQYSYSFSKEAKYQVL 360
 Db 345 TGWYGRPRSEPHFTLDGNSFPYKLTISNEYKEMPGGRNLYKIQSLDYTKTCLSCPBLNPQCQYSYSFSKEAKYQVL 404
 Db 361 TSDYLYTISNEYKGMPGGRNLYKIQSLDYTKTCLSCPBLNPQCQYSYSFSKEAKYQVL 420
 Db 405 TSDYLYTISNEYKGMPGGRNLYKIQSLDYTKTCLSCPBLNPQCQYSYSFSKEAKYQVL 464
 Db 421 RCSPGGLPLYTHS SYNDKGLRYLEDSALDKMQLQVYOMPSKCKLDFILNETKFWYQML 480
 Db 465 GCMPGGLPLYTHS SYNDKGLRYLEDSALDKMQLQVYOMPSKCKLDFILNETKFWYQML 524
 Db 481 PPEFDKSKCKYKPYLDDIVYAGPCSKOKADTYVFLNWATYLASTENIYASSTAHQHITYSHLHQCPSL 540
 Db 525 PPFDKSKCKYKPYLDDIVYAGPCSKOKADTYVFLNWATYLASTENIYASSTAHQHITYSHLHQCPSL 584

RESULT 3

S23752 dipeptidyl-peptidase IV (EC 3.4.14.5) alpha chain - mouse
 N;Alternate names: CD26 alpha subunit; THM alpha subunit
 C;Species: Mus musculus (house mouse)
 C;Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 10-Sep-1999
 C;Accession: S23732; A46465; A563030
 R;Marguet, D.; Bernard, A.M.; Vivier, I.; Darmoul, D.; Naquet, P.; Pierres, M.
 J. Biol. Chem. 267, 220-228, 1992
 A;Title: cDNA cloning for mouse thymocyte-activating molecule. A multifunctional ecto-di-

Qy	541 MHA INRRLGTFEVEDQI EEARQFESXKMGFTVDNKRIAWGM SYGGVTVTSVNLGS GSGGVFKCG 600	Db	522 MLPPQFDRSKCYPLLQQVYGGPCSQSVRSVPAVNWSYLASKEGMVIALVDRGTAFAQG 581
Db	585 MHA INRRLGTFEVEDQI EEARQFESXKMGFTVDNKRIAWGM SYGGVTVTSVNLGS GSGGVFKCG 644	Qy	538 DKIMHAINRRLGTFEVEDQI EEARQFESXKMGFTVDNKRIAWGM SYGGVTVTSVNLGS GSGGVFKCG 587
Qy	601 IAVAPVSRMVEYYDSVTEYRVMGLPTPENLDLHDYRNSTMVSRAENFQVKOBYELLIHGTA DN 660	Db	582 DKLUYATYTRKLGYYEVEDQTAVRKFLEMGFDEKRIAWGM SYEIRFITGPCIWNMSFQ 641
Db	645 IAVAPVSRMVEYYDSVTEYRVMGLPTPENLDLHDYRNSTMVSRAENFQVKOBYELLIHGTA DN 704	Qy	588 NYLGSGSGGVFKCGIAVAPVSRMVEYYDSVTEYRVMGLPTPENLDLHDYRNSTMVSRAENFQK 647
Qy	661 VHFQOSAQISKVALDVGDFQAMMFTDEDHGIASSTAHQIYTHMSHIKOCFSI 715	Db	642 MWSGGSSL-QGILRLVCLHR ----- BIHGSPIKDNLEHYKNSTMARAEFRN 690
Db	705 VHFQOSAQISKVALDVGDFQAMMFTDEDHGIASSTAHQIYSHMSHFLQQCFSI 759	Qy	648 VEYLILINGTADDNVHFOQSAQISKVALDVGDFQAMMFTDEDHGIASSTAHQIYTHMSH 707
RESULT 4			
I38593	Fibroblast activation protein-alpha - human	Qy	691 VDYLHLHGTDADDNVHFOQNSAQIQLAKLYNAQVDFQAMMFTDEDHGIASSTAHQIYTHMSH 749
N;Alternate names: FAP-alpha	C;Species: Homo sapiens (man)	Db	708 FIKQCFSL 715
C;Date: 09-Mar-1996 #sequence_revision 09-Mar-1996 #text_change 29-Aug-1997	C;Accession: I38593	Db	750 FIKQCFSL 757
Riscanin, M.J.; Raj, B.; Calvo, B.; Garin-Chesa, P.; Sanz-Morcasi, M.P.; Healey, J.; Ol	A;Reference number: I38593; MUID:94261645; PMID:7911242	Qy	768 FIKQCFSL 715
Proc. Natl. Acad. Sci. U.S.A. 91, 5657-5661, 1994	A;Accession: I38593	Db	768 FIKQCFSL 757
A;Title: Molecular cloning of fibroblast activation protein alpha, a member of the serin	A;Status: preliminary	Qy	768 FIKQCFSL 757
A;Residues: 1-759 <RES>	A;Molecule type: mRNA	Db	768 FIKQCFSL 757
A;Cross-references: UNIPARC:UPI0000175B7D; EMBL:U09278; PID:9507750	A;Cross-references: UNIPARC:UPI00016A7F9; GB:9507750	Qy	768 FIKQCFSL 757
C;Genetics:	A;Status: preliminary	Db	768 FIKQCFSL 757
A;Gene: GDB: FAP	A;Residues: 1-803 <RES>	Qy	768 FIKQCFSL 757
A;Map position: 2q23-q23	C;Superfamily: dipeptidyl-peptidase IV	Db	768 FIKQCFSL 757
C;Keywords: transmembrane protein	C;Keywords: transmembrane protein	Qy	768 FIKQCFSL 757
Query Match			
Best Local Similarity 49.8%; Score 1930.5; DB 2; Length 759;	Matches 388; Conservative 134; Mismatches 201; Indels 35; Gaps 11;	Qy	768 FIKQCFSL 757
Best Local Similarity 49.2%; Score 1930.5; DB 2; Length 803;	Matches 388; Conservative 134; Mismatches 201; Indels 35; Gaps 11;	Qy	768 FIKQCFSL 757
Matches 388; Conservative 134; Mismatches 201; Indels 35; Gaps 11;	Qy	768 FIKQCFSL 757	
2 YTKRLYSLRWHISHEYLYKQ-ENNIVLNVAEYGNSSVFLENSTFDEFGHSIN-DYSIS 58	Db	768 FIKQCFSL 757	
51 TPSYKTFPNWISQEQYHQSAADNNIVLNIEGTQSYTILSNRTM---KSVNASNYGLS 106	Qy	768 FIKQCFSL 757	
59 PDGQFILLEYNNTQWRSRHTASYDIDLNKRLQITERIPIINTQWTSPVGHKLAYW 118	Db	768 FIKQCFSL 757	
107 PDRGEVYLSDYSKLRWSRHTASYDIDLNGEVRGNBLPRIQYLWSPVGSKLLAYV 166	Qy	768 FIKQCFSL 757	
119 NNDIYVKCBLPNLPSYRITMTGEDIILYNGITDMYEEEVFSAYSALWISPGNTFLAYQF 178	Db	768 FIKQCFSL 757	
167 QNNYLKQDGPDPFQITFNGREKINGKIPIDPYVYEMPLPKYALWISPNCKFLAYEF 226	Qy	768 FIKQCFSL 757	
179 NDTPEVPLLEYSFYSDESQYQPKTIVRPYPKAGANPTKFFVNTDSIISVNTNATSIQT 238	Db	768 FIKQCFSL 757	
227 NDKDIPVAYASYCD--QYPRTNINIPPKAGAKNPVTRIFIDTYFAYGPQ-- EVP 281	Qy	768 FIKQCFSL 757	
239 APASMLIGDHYLCDVTTWATQERISLQWTLRRRIQYNSYMDICDYDESSGWNLCLVARQHJEM 298	Db	768 FIKQCFSL 757	
282 VPAMIASSYFSWMLTWWTDERVLQWNRVQNVSVLICDFRDWQWDCPATOEEIE 341	Qy	768 FIKQCFSL 757	
299 STICWUGRRPSEBFHTLDGNSFYKIIKISNEEGYRHICFQIDKKDCTPIKGTMWEVIGIE 358	Db	768 FIKQCFSL 757	
342 SRTSWAGGFVSRFESDAISYKIFPDGDYKHAIIKDVTEAAGITQITSGKWEANIF 401	Qy	768 FIKQCFSL 757	
359 ALTSDFLYYISNEYKGMPPGRNLQYKIQDSDY-TKVTCIUSCEINPERCCOYYSVSFSKEAK 417	Db	768 FIKQCFSL 757	
402 RVTQDSLWSSNEEEYFCCRNYIYRISGSYPPSKKCVTCHLRKERCCYTSFSDYAKY 461	Qy	768 FIKQCFSL 757	
418 YOLRCSGPGLPLYLHSSYNDKGRVUELDNSALDKMUNVOMPSKULDFILNETKFWQ 477	Db	768 FIKQCFSL 757	
462 YALVCYGGCIPISTLHDGRTDQEKLILNKAQNLQPLREEKIKLVEDEITLYWK 521	Qy	768 FIKQCFSL 757	
478 MILPPHEFDKSCKYPLLLDLYAGCPCSQKADTVERLNWATYLASTENIIVASFDGRGSGYQ 537	Db	768 FIKQCFSL 757	
Query Match			
Best Local Similarity 33.2%; Score 1117; DB 2; Length 803;	Matches 241; Conservative 150; Mismatches 290; Indels 44; Gaps 20;	Qy	768 FIKQCFSL 757
Best Local Similarity 33.2%; Score 1117; DB 2; Length 803;	Matches 241; Conservative 150; Mismatches 290; Indels 44; Gaps 20;	Qy	768 FIKQCFSL 757
70 YVKQRHHSYTASYDIDL-NKRLQITEERPN-INTQWTSPVGHKLAYWANDIYWK 126	Db	768 FIKQCFSL 757	
145 VEPYQHSYTGTYVLSKIPHGPDPQSLDPPEVSNAKLUQYAGNGCPKGQQLIFIFENNIYCA 204	Qy	768 FIKQCFSL 757	
127 EPNLPSYRITVGYKEDDIILYNGITDWYEEEVFSAYSALWISPGNTFLAYAQNNDTEVPL 186	Db	768 FIKQCFSL 757	
205 HVGKQAIRVVTGKEGTYINGLSDWLYEEELKTHAHHWSPDGTRLAYAINDSRPIM 264	Qy	768 FIKQCFSL 757	
187 EYSFYSDESLOQPKTIVRPYPKAGANPTKFFVNTDSIISVNTNATSIQTAPASMLIG 246	Db	768 FIKQCFSL 757	
265 ELPTYTGSI--YPTPKRYHYPKAGSENPSISLHV ---GLNGPTH--DLEMMPDDPRMR 317	Qy	768 FIKQCFSL 757	
247 DHYLCDVTTWATQERISLQWTLRRRIQYNSYMDICDYDESSGWNLCLVARQHJEMSTGNGVR 306	Db	768 FIKQCFSL 757	
318 EYITMYKWAITSKAVTWNRAQNTSILTC--DATTG---VCTKXKH-EDBSEAWLHR 370	Qy	768 FIKQCFSL 757	
307 FRPSEPHFTLDGNSFYKIIKISNEEGYRHICFQIDKKDCTPIKGTMWEVIGIE 358	Db	768 FIKQCFSL 757	
371 -ONEEPFSKDCDKFFTIRAPQGGRGKFYHITVSSQNPSSNDNQISITSGDWDVTKIL 429	Qy	768 FIKQCFSL 757	
359 AL-TSDPYLYISNEYKGMPPGRNLQYKIQDSDY-TKVTCIUSCEINPERCCOYYSVSFSKEAK 416	Db	768 FIKQCFSL 757	
430 AYDEKGKNIYFLISTS--DLPLRRRQLQYSANTEGNFNRQCLSCDL-VENCTYFSASFHSMSD 486	Qy	768 FIKQCFSL 757	
417 YYQLRCGPGPLPLYLHSSYNDKGRVUELDNSALDKMUNVOMPSKULDFILNETKFWY 476	Db	768 FIKQCFSL 757	
487 FPLLKCBGPGVPMVTVHNTDOKKMPFDELETNEHVKAINDQMPKVEYRDIIDYNLPM 546	Qy	768 FIKQCFSL 757	

477	QMLPPHFDSSKCPULLDYGAGCOSQAKADTVERLNWATYLASTENITIVASFDORGSGYQ	536
547	QILRKPATFTTTHPLLVDTGQSOSVAKEFVSWETMVMSHGAVVVKCDORGSSFQ	606
537	GDKMHAIRNRLGTFEVEDOJEARQFSKMGFDVNKRIAIWGSYGGYTSMYL--GSG	593
607	GTLKLHEVRRLGLLESDQMEAFTMLKEQYIDTRYAVFGDYGLSTYLPAKGEN	666
594	SG-YFKCGIAVAPYSRWEYDSVYTERYMGPLPTBENDLHD--YRNSTVMSRAENKQEY	650
667	QGQIFTCGSALSPITDFKLYASAFTSERVIGL---HGIDLNRAYEMTKVARYSALEEQF	722
651	LJIITGADDNVHFQOSAQIISKALVDGVDFQAMWYTDDEGHGIASSTAHQIYTHMSHPK	710
723	LIIHPTADEKIHFOHTAELJTOLIRGKANYSLQTYPDESHEYFTSSSLKQHLVRSIINFV	782
711	QCFL 715	
783	ECFR1 797	

RESULT 6

53331 aminopeptidase like protein - human
 species: Homo sapiens (man)
 date: 02-Jul-1996 #sequence_revision 02-Jul-1996
 accession: U154331
 ;Tokonami, N.; Doi, K.; Wenthold, R.J.; Wada, K.
 ;Mol. Genet. 2, 1037-1039, 1993
 Title: Non-conservation of a catalytic residue in
 Reference number: U154331; MUID:93172805; PMID:81131
 Accession: U154331
 Status: preliminary; translated from GB/EMBL/DDB

Molecule type: mRNA
 Residues: 1-865 <RES>
 Cross-references: UNIPROT:P42658; UNIPARC:UPI0000047378; GB:MG96859; NID:g306705; PIDN:
 Superfamily: dipeptidyl-peptidase IV

Query	Match	Length
Best	28.8%	Score 1117; DB 2;
Local Similarity	33.2%	Pred. No. 3.3e-60;

Matches	Conservative	Mismatches	290:	Indels	44:	Gaps	20:
241:	/	/					
11	RWISDHSLVYK-QENNILVNAEYNSSVLFENSTPDEFGHSINDYNSISPDCQFILLBYN	69					
149	KWISSTEFTIREQKGTVRLLNVNTSTVLIEGKIESL-RAIR-YEISPREYALFSYN	206					
70	YVKQRWHTSYSDYDYL--NKROJ-LITEER-TPN-NTQWTSVSPVGHLKAYWNNDIYVKI	126					
207	VEPIQHSTGYVVISKI PHDDPQSJDPPVSNALKUQAGNGPKGOOLIFPNENIYCA	266					
127	EPNLPSYRITWTGKEDDLYINGITDWYTYEEVPSAYSALWWSBNGTFLAYAQENDTEPLI	186					
267	HVGKQAIRVYSTKGKESVYINCLSDLUYEEELLKTHIAHWNSPDGTRLAYA1INDSRPIM	326					
187	EYSFFSDESLOQPKTVRVPYKAGA\NPTVKEFVWNTDSLSSTVNATNSQIQITAPASMLIG	246					
327	ELPTVGTSSI--YPTVKPYHYPKAEGSENPSISLHV1 --GLNGPTH--DLEMMPDDPRMR	379					
247	DHYLQDVTWATOERSLQWLRRIQNYSMIDCYDSSGRNCLVAROHIEMTGTGWGR	306					
380	EYYTVMVKWATSTKAVTWNRAQNSVILTC--DATG--VCYTKH--EDBESEAWLHR	432					
307	FRPSPHFTLDGNSFPKLIISNEGGHRHICYF-----Q1DKKDCTPITGTWEVIGIE	358					
433	-QNEBPVFSKDGKFFFIRAIPOGGRGKFHTIVSSQNPNSNDNQISITSGDWDYTKIL	491					
359	AL--TSDLYIYISBYKGMPGGRNLYKIQLSDTYKVTCSCLENPERCQYYSVSFSKEAK	416					
492	AYDEKGKNKIVFLSTE -DLPLRRQLYSANTEGNFNRQLSCDL-VENCTYFSASFHSMD	548					
417	YYQFLRCGSGPUPLYTHSYNDKGRLVEONSAIDMQLONTQMPSKLDFIILNEKFWY	476					
549	FEFLKCEGCPGPVPMYVHNTTDDKKKOFDLENEHVKAINDROMPKYERYDIEDDNLPMLP	608					

RESULT 7
 A41793
 dipeptidyl aminopeptidase-like protein (EC 3.4.14.-) short form, DPPX-S - bovine
 C;Species: Bos primigenius taurus (cattle)
 C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
 C;Accession: A41793
 C;Reference number: A41793 ; MUID:92108018 ; PMID:1729689
 A;Title: Differential expression of two distinct forms of mRNA encoding members of
 A;Status: Preliminary; not compared with conceptual translation
 A;Molecule type: mRNA
 A;Residues: 1-803 <WAD>
 A;Cross-references: UNIPROT:P42659; UNIPARC:UPI000002A81C; GB:MT6429; NID:g408719
 A;Note: sequence extracted from NCBI backbone (NCBIP:75138)
 C;SuperFamily: dipeptidyl-peptidase IV
 C;Keywords: dipeptidylpeptidase hydrolase, Glycoprotein
 F;27,34/Binding site: carbohydrate (Asn) (covalent) #status predicted
 Query Match Score 1094 ; DB 2 ; Length 803 ;
 Best Local Similarity 32.9% ; Pred. No. 1.2e-67 ;

Matches	Conservative	153;	Mismatches	287;	Indels	48;	Gaps	22;
Qy	11	RWISDEHYLK-QENNILLVFNAAEYGGSSVFLENSTDFEGHS	INDYSIS	SPDGQPLILEYN	69			
	:	:	:	:	:			
Db	87	KWISDXKEFIREQKGSSVILRNVEVTNTSVLIGKKIESL-RAIR-YEISPDREALPSYN	144					
Qy	70	YVK0K9HSTSYASDIDL- -NKROLITTEERION-NTORIVTWSWPVGKCLAYWNNNIIYVKKI	126					
	: :	: :	: :	: :	: :			
Db	145	VEPI0CHSYSTGYYVLSKIPHDGPQSDDPPEVNAKLQTAGWPKQKQLIFENNNIYCA	204					
Qy	127	EPNLSSYRITWGTGKEDILYNGLTDWYEEEVFSAYSALWWSPNGTFLAYAQAFNDTEPVPLI	186					
	:	:	:	:	:			
Db	205	HVGKQARVYSTGKEGVYINGLSDWLYEEELKTHIAHWWSSPDTGLAYATINDSRVPVM	264					
Qy	187	EYSFYKDESLOYPKTRVPKAGANAPTKREFVNTDSSLSSVTNATSQIQTAPASMLIG	246					
	:	:	:	:	:			
Db	265	ELPTTIGS- -VYPTAKPHYKAGCNPMSISLHV1--GLNGPTH--DLEMTPDPDRMR	317					
Qy	247	DHYLCIWTQTLWERTQERILQWLWERTQYIYDDESGRWNLLVARTQHJEMSTGJWGR	306					
	: :	: :	: :	: :	: :			
Db	318	EYITIVKWRKATSTKVAVNWLSRAQNSISLTC-DATIG-- -VCTKRGH-EDBSBWLHR	370					
Qy	307	FRPSEPHFTLDGNSPFIKISNEEG---YRHICYP---QIDKDKDCTPITKGTMWEVIGTE	358					
	:	:	:	:	:			
Db	371	-QNEBVFSSDGRKFVFRAIPOGGQGRFHYHTVSSSPNNSDNIQSITSGRDVTK1-428						
Qy	359	ALTSD---YIYIISNEYKGMPGGRNLYKIQSLDYTKTCLSCELNPERCQYSVSFSKE	414					
	:	:	:	:	:			
Db	429	-LSYDEKRSSVFLSTYE- -DUPRRRQOLYASSTYGSFVRQCLSDF-VDNCTYFASPSPG	484					

Qy	415 AKYQLRCSCPGPGLPLYLHSSVNDKGLRVLDENSALDDKMLQNYOMPSTKLLDFIILNETKF	474	Qy	338 QIDKRDCTFITKGTVWYIG-----TEALTSDFLYYSNEYKGMPGGRN	380
Db	485 ADFFLKCGPGPVRSVANTDKRMFDLMEHVAISDQMPKBYRKLETDDNL	544	Db	405 PFGNGSPPIYLTSQAWDVTGPAHIDGDFGVYFLATLKDSTERHLVVS-----	453
Qy	475 WYQMLLPHEFDKSKKYPILLDYYAGPCGQADTYFRLNWATYLASTENILIVASFDGRSG	534	Qy	381 LYKIQLSDTYTKTCLSCELNPERCQYQSVSFSEKAKEYQLRCSGPGLPLYLTHSVDNG	440
Db	545 PIQILKPAFTDTDAHYPLLUVDGTPGGSVAERFAVTTWETMVSSHGAVVYKCDGRSG	604	Db	454 LDTLEIYGTD-----NGEDEGYSTSFSFSPFGDFTVLYNGHPDPWOBLSRKDKDY	505
Qy	535 YOGDKIMMAINIRRIGTFTFEVDQEAAQFQSMGEFDNRKRIATGWGTYGGVTSMVL--G	591	Qy	441 LRVLDENSALDGMQLNQVOMPSKSLKLDFILNETKFWOMILPPFDIKSKYPLLDVYAGP	500
Db	605 FGCFRLHMEVRRRGLSLEEKDOMEAVRYMKEPIKDRVAVGKDYCGYLSTLPPKG	664	Db	506 CLSLETSRSLRQOLQSSLTILPSVEYGLITENDTMRPRNEDVNCYKPLFPAYGGP	565
Qy	592 SGSG-VFKCGIAVAPSRMEYDSVTFYRMLGPTPEONLHD--YRNSTMSPRAENPKQV	648	Qy	501 CSOKADTYFRLNWATYLASTENILIVASFDGRGSGYQGDKLMHAIRRLGTFEVEDQIE	558
Db	665 DQGAPVFSCGSALSPITDPLKYSASFSPRYGL---HGLDNAYEMXKAHRVSALEGQ	720	Db	566 GSQQAKLFERDVEQAYLAHPDFFIVTLDGIGNGNAFRYSVSRHLGGEWEVSxDQG	625
Qy	649 EYLLHGTAADDNVHFQOSAQISKALVYDGVDFQAMWYTDDEDHGIASSATAHOIHYTMNSHF	708	Qy	559 AARQFSKMGFYDNCRIAIWNGSYGGYTTSMYLGSGSGVFKCGIAVAPVSRMEYDSVYTE	618
Db	721 QFLYTHATADEKTHFQHTPAELITOLIKGRANYSLSQIYDSEHYFFSSALQOHLHRSTIGF	780	Db	626 AGRFWADLPLPFDENHYGIWNGSYGGYGLTLKTLET-QDVSYSGMAVAPTDWRLYDSVYTE	684
Qy	709 IKQCFSL	715	Qy	619 RYMGLPPEPDNLHDYRNSTMSRAENPKQVE-YLLHGTADDNVHFQOSAQISKAL---	673
Db	781 FVECFRI	787	Db	685 RYMDL--PYKNEYGGKNSQIH-DYKEFKQLKRFFVAHGTDDNNFHQSMLMDGILNLAN	741
Qy	RESULT 8		Qy	674 ---VDVGVDFOAMWYTDBDHGTASSTAHQHITYTHMSHFI	709
Db	TA1703 d-peptidyl aminopeptidase - fission yeast (Schizosaccharomyces pombe)		Db	742 CYNQDMAV-----FPDSAHSISYANASLSTYHRLSEWI	774
Qy	Species: Schizosaccharomyces pombe		RESULT 9		
C;Species:			A30107	dipeptidyl aminopeptidase B (EC 3.4.14.-) - yeast (Saccharomyces cerevisiae)	
C;Date:	03-Dec-1999 #sequence_revision 03-Dec-1999	#text_change 09-Jul-2004	N;Alternate names: protein YHR028C		
R;Murphy, L. : Harris, D. : Barrell, B.G. : Rajandream, M.A. ; Wood, V.			C;Species: Saccharomyces cerevisiae		
C;Accession:	T41703		C;Date: 07-Jun-1990 #sequence_revision 30-May-1997 #text_change 09-Jul-2004		
A;Reference number: Z22011			C;Accession: S46780; A30107		
A;Status: preliminary ; translated from GB/EMBL/DBJ			R;Du, Z.		
A;Molecule type: DNA			A;Submitted to the EMBL Data Library, June 1994		
A;Residues: 1-793 <MUR>			A;Description: The sequence of <i>S. cerevisiae</i> cosmid 8082.		
A;Cross-references: UNIPROT:O14073; UNIPARC:UPI000013AB95; EMBL:AL031180; PIDN: CRA20138.			A;Accession number: S46773		
A;Experimental source: strain 972h-; cosmid c2E11 -chimeric			A;Molecule type: DNA		
C;Genetics:			A;Residues: 1-818 <DUZ>		
A;Gene: SPAC2B11.08			A;Cross-references: UNIPROT: P18962; UNIPARC:UPI0000031A5F; EMBL:U10399; PID:9500689; PIR:R;Robert's, C.J.; Pohlig, G.; Rothman, J.H.; Stevens, T.H.		
A;Map position: 1			J. Cell. Biol. 108, 1363-1373, 1989		
C;Superfamily: dipeptidyl-peptidase IV			A;Title: Structure, biosynthesis, and localization of dipeptidyl aminopeptidase B, an iron-binding protein. The authors translated the codon ACC for residue 572 as Asn		
Query Match 24.2% Score 940; DB 2; Length 793;			A;Reference number: A30107; MUID:89174971; PMID:2647766		
Best Local Similarity 31.1% Pred. No. 5 3e-57;			C;Genetics:		
Matches 236; Conservative 133; Mismatches 260; Indels 122; Gaps 26;			A;Molecule type: DNA		
Qy	20 YKQNNI-----LVFNAYEG-----NSSVF-LENSTFDEFQHSTINDYSTSPDG	61	A;Cross-references: 1-818, 'H', 94-124, 'N', 126-181, 'LRLLET', 189-199, 'N', 201-365, 'DFKRGKERKF', 376-57		
Db	69 YKQVNWIDNSQGLKDTRFLXGDLNIQDYNLNKTFLVSDFLYNGTSDSYSFDA	128	R;Cross-references: UNIPARC:UPI000172A3C; EMBL:X15484		
Qy	62 QFILLENTYVKQWRHSYTSASYDYLINRQLITEERIIN-----NTQW---VTVWSPVGH	112	A;Note: The authors translated the codon ACC for residue 572 as Asn		
Db	129 KRYLVSVNSQQRHSSPQQYLYN-----TETKDNMLGQDNNEHTITLAEWSVGH	181	C;Keywords: dipeptidyl-peptidase IV		
Qy	113 KLAYWNNDIYV-KIEPNLPSYTRITWTGEDIYNGITDWWYBEVEFAYSALWWSPNGT	171	P;30-55/Domain: transmembrane #status predicted <TMM>		
Db	182 QLSFVYNDLIVRKNDGNV--QRITYDGTVD-VFNGLTDWIBEEVLSPSTIWWSPSD	238	F;63,79,110.139,392,421/Binding site: carbohydrate (Asn) (covalent) #status predicted		
Qy	172 ELAYAQFDTEPVLIJEYFSYSD----SLOQPKTVRVPYPPAKAVNPTVKEFVNTDS	225	Query Match 23.4% Score 908; DB 1; Length 818;		
Db	239 KIAFLKLNSEIPTHYPLXTBDELPSPEFDNKDMAIKYKPGPNSVSLFV--ADL	296	Best Local Similarity 32.6% Pred. No. 9.1e-55;		
Qy	226 LSSTVNATSIQTAPASLIGDHYLCDTWATQERISQWLRIQNYSMDCYDESSG	285	Mismatches 251; Conservative 119; Indels 92; Gaps 27;		
Db	297 NSNASSFLRSLRATDSDHYLYKQBNL-----VFNABYGN-----SVYLENSTPDE	351	Qy	1 NTYRLKUYSLRATDSDHYLYKQBNL-----VFNABYGN-----SVYLENSTPDE	47
Qy	286 RWNLVQHRIEMSTTGWGRFRESEPHFTL-----GNSFYKLISNEESYRHTCYF	337	Db	80 NTFQPKYHQLNISDN---KIESNDLGLYXYTFMNDSYVKSVDYDSYNSVLEGKTFH	135
Db	352 ----IHTVKTECLEGGY-EVQOSAKMFPLNNSLVWENWSGQYFDILALDD-YNHLPFI	404	Qy	48 FGHSINDSYI--SPDGQFILLBEYLYNQKQWRHSYTSASYDIOLNKRQLIITEERIPNNTQWV	105

Db 136 NGQNLTVESTASPDLKRLJIRTNSYQNRHSTFGSYFVYDKSSSF--EEIGNEVALA 192
 Qy ::::: :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 106 TWSPVGHKLAVWNNNDIYVKEIPNLESYRITWTGKEDIDYNGITDLYVKEIPNLESYRITWTGKEDIDYNG 147
 Db 193 IWSPNNTDIAYQDNNIYIYSAISKTKTIRAVTNDSSFLENGKPDWVYEEBFEDDKAAW 252
 Qy ::||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 166 WSPNGTFLAYAQFNDETEVP-LIEYSFSDESLOYPKTYRVPPKAGAVNPTVKEFFVNT 223
 Db 253 WSPTGDLAFURKIDESBVEPIPY-YVQDBEKD1PEMRSIKYPSGTGPRAHAWVY-- 309
 Qy ::||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 224 DLSSTVTNTSQTAPASMLIGDHYLCDVYTWAQRERISLQWLRRIQNYSYMDCDYES 283
 Db 310 ----SMKGRSFPHRISGNRKDGSLSLITEVWVGNGNLVKTDRSSDLTVEFLDTAK 365
 Qy ::||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 284 SGRWNCLVAQHIMSTGW---VGRFRPSEPHFTLDNSFSYKIIISMEGYRHICYFQ 338
 Db 366 TSN----VVVN-ESSNGGWWEITINTLFPANETFDPRHNGYDILP-IGGYNHLLAPE 418
 Qy ::||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 339 -IDKEDCTFITKGTMWVI-GEAIALTS-DYLYTANISNEYKCMPGGRNLKYIQL---SDYTKT 391
 Db 419 NSNSSHYKTILEGKRNBLVNGPLAPSMDENLYFISTRSSSTE-RHVTYDLSRSPNELL 476
 Qy ::||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 392 VTCLSCELNPERCQYVSFSKEAKYQQLRCSCGP LP--LYTHS-----SVNDK 439
 Db 477 VTDTSBD-----GYVDVSFSGRRFGFLTYKGPKVYQKIVDPISRKAECDKDGNVLGK 530
 Qy ::||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 440 GLRIVEDNSALDKMLQNVQMPSKKUDFLINNETKWM-----ILPPIHFDK-SKVYP 491
 Db 531 SLYHLEKNEVLTKILEDYAPRKSFQSKADTFVRLNWATYLASTENITYASFDRGSGYQGDKLMHAIRRQCTF 551
 Qy ::||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 492 LLLDYVYAGPCQCSOKADTFVRLNWATYLASTENITYASFDRGSGYQGDKLMHAIRRQCTF 551
 Db 591 VFFFAYGPNQSQVVKTFSTGFNEVVAASQNLAIYVVDGRTGPKQDFRSYVDRD lady 650
 Qy ::||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 552 EVEDQIBAAQFQSKNGPVDNKRIAIWGSYGGYTTSMVLGSGSG-VRKGIAVAPVSRMK 610
 Db 651 EARDQSAASLYGLSTFVDQKISLFGWSYGGYLTLKTLEKDGRHRFYKGMSVAPYDTRKR 710
 Qy ::||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 611 YDSVYTERVNGLPPPEDNULDHY----RNSTVNSRAENFKQVSYLLIHTGADDNVHFQ 665
 Db 711 FYDSVYTERVM-HPQENFDGYESSVNTALAQNRF---LLHGTGDDNVHFQN 763
 Qy ::||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 666 SAQISKALYDGVDF-QAQWYTDDEDHGIASSTAHOIYTHMSHPIKOCF 713
 Db 764 SLKFLDLDLNGVENDVYDVAVFPDSDHSTIRYHNANVTFDQLDWKRAF 812

RESULT 11

T19514 hypothetical protein C27C12.7 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #Sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T19514
 R:Thomas, K.
 C:Reference number: Z19134
 A:Accession: T19514
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Genetics:
 A:Residues: 1-829 <WIL>
 A:Cross-references: UNIPROT:Q18253; UNIPARC:UPI000007DC53; EMBL:Z69883; PIDN:CAA93743.1
 A:Experimental source: clone C27C12
 C:Genetics:
 A:Gene: CBSP:C27C12.7
 A:Map position: X
 A:Intron: 4/3: 51/1: 123/3: 166/3: 188/2: 279/2: 392/3: 501/3: 583/1: 606/2: 64
 C:Superfamily: dipeptidyl-peptidase IV

Query Match 19.2% Score 745; DB 2; Length 829;
 Best Local Similarity 28.8%; Pred. No. 1..8e-4;
 Matches 227; Conservative 128; Mismatches 282; Indels 150; Gaps 35;

Qy 8 YSLRWDHBYLYKQENNIVLNFVNAEYGNSSVPLENSTPDEFGHSINDYSISPPDQFLU-- 65
 Db 105 YDTIWLPGDFSVQMMND---FTIRKQMKKIPFQSSVAEPFFN-----NEYVKAL 151

RESULT 10
 S66261 x-Pro dipeptidyl-peptidase (EC 3.4.14.11) - Flavobacterium meningosepticum
 C:Species: Flavobacterium meningosepticum
 C:Accession: S66261
 C:Date: 28-Oct-1996 #Sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
 C:Accession: S66261
 R:Kabashima, T.; Yoshida, T.; Ito, K.; Yoshimoto, T.
 Arch. Biophys. Biophys. 320, 123-128, 1995
 A:Title: Cloning, sequencing, and expression of the dipeptidyl peptidase IV gene from *F*
 A:Reference number: S66261; MUID:95314307; PMID:7793370
 A:Accession: S66261
 A:Status: preliminary
 A:Molecule type: DNA
 C:Superfamily: dipeptidyl-peptidase IV
 A:Cross-references: UNIPROT:Q47900; UNIPARC:UPI00000B2115; EMBL:D42121; NID:9577283; PID:9577283
 C:Keywords: dipeptidyl-peptide hydrolase

Query Match 20.4% Score 791.5; DB 2; Length 711;
 Best Local Similarity 30.3%; Pred. No. 8..6e-47;
 Matches 216; Conservative 125; Mismatches 274; Indels 97; Gaps 24;

Qy 31 AEYGNSSVFLENSIFDEFGRHSINDYSISPDGQFILLETNYVKQWRHSYSTASDYLINKR 90
 Db 60 AKTYSKTSOREKNTIVD--GSFOGXTFSNDESKILLQKSOSYRHSPLGKEPVKDKR 116

QY 66 ---LEYNYKQ-----WHRSYTASYDIVDINKROLITBERIPNTOW-----v 105
 Db 152 SSNNRKYAYGSKKVNEMLRISAEYLHYRKINNKTVTSE-----QHVGPEEINSLQDAF 204
 QY 106 TWSP---VGHKLAYWNNDIYVKKIEPNLPSYRITWT-GKEDILYNGITDWWYEEBEVSYS 162
 Db 205 YNNPNSNDSDEVVYHNLYYQNDPEKEPDGAIQ.TVGGSTENRGFLAWMLYEBEILLESS 264
 QY 163 ALWSPNGTFPLAYAQNDTEPVLTFSDESQYPTKVRVPYPKAGAVNPV-VKFVV 221
 Db 265 AVWTPSPGRVSYLRFDPDREVN.FLPYTYDD-SYVEYFELPPKAGYQNNTLVTQIW 323
 QY 222 NTDSLSSVTNATSTOITAAPASMLIGDHTXLCDDVTWAT-----QERTISLQWIRRICNY 272
 Db 324 DSENHKIVTAPPNLSAAN---GDIYVLTNWKWITMPRNGSDLGEERLVTVWANRDQNH 379
 QY 273 SVMJDICDYDESSGRWNCLVA-----ROHIEMSTTGNWGRFRPSE-----PHPFILD 317
 Db 380 VYFSUCNEQD-----CNALSFQSFSIDNRLQWVSKPKDVKVGP-PTETGFLTVLPHKHDD 432
 QY 318 GNSPYKISNEEGYRHYCIVFQIDRKDCTPITKGTMW-----VIG----IEALT-----361
 Db 433 GNI-----YNHVARVYELDGTGKTKWGENFDVILVGYSSKIDALITFSAWGD 482
 QY 362 --SDYLYXISNEYKGMPGGRNLNYKIQLSDYTKVTCLSCELNPERCQYY-SVFSKEXAKYY 418
 Db 483 GVGFBSTYVRE-----AMTSNKKTQLQVKTD-----QFEBCKTQSQSADPTGCRRI 529
 QY 419 QLRSGSP--GLPLYLTHSSWYNDKGLRLEDNSALDKMLQNQVMQPSKCKLDFILL-NETKFW 475
 Db 530 VVQCBKPFDNTRLYV--DyVDTDTKIMLEGGT--KAVIPFDVFNMKGLKLPSGIDGH 585
 QY 476 YQMLPPHPDPSKRYPLJUDVYAPCS---QKADTVPRLNATYLASTENTIVASFGGR 531
 Db 586 YMMLTSPANLGDAKIPLLDIYGPDSKOVFQKPTPAK---AIIQIVSQYDIAYARDV 641
 QY 532 GSGYQGDKIMHAINRRLGTFFEVEQTEAAROF-SIKMGFYDNKRTIAIWWSYGGVTSWV 590
 Db 642 GTGRGFWDWEAVYRKLGAEVWDTDLMIRAFINTFGTIDEIAWMWSYGGFLTSIA 701
 QY 591 GSGSG--VFKCGIAVAPVSRWEYDSVTERYMLGLPTPENLHYRNSTMMSRAENPKVE 649
 Db 702 IKDQGELVRCAISIAPVTDKYYDSAYTERYLG-QPARENQYINTNNTVIPHARNNTVK 759
 QY 650 YLLHGTAADDNVHFFQSAQISKAULDVGYDFQAMWYTEDDHGIAASSTHQ--HIYTASH 707
 Db 760 YLLAHGERDDNVHYQNSARWSEALQONGTHFTOLVYANBAH---SLSHKLFLHYGEVQR 815
 QY 708 FI-KQCF 713
 Db 816 FLMNDCF 822

RESULT 1.2

A4937 dipeptidyl aminopeptidase (EC 3.4.14.-) - yeast (Saccharomyces cerevisiae)
 N; Alternate names: protein 05045; protein YOR219c; protein YOR50-9
 C; Species: Saccharomyces cerevisiae
 C; Accession: A4937; S45451; S63046; S67112; S71721
 R; Santa Anna,A., S.; Herskowitz, I.
 Submitted to the Protein Sequence Database, July 1993
 A; Reference number: A49/37

A4937 aminopeptidase; UniProt:P331894; UNIPARC:UPI0000136060; GB:L21944; NID:g347196; PIDN:
 R; Anna-Arruda, S.S.; Herskowitz, I.
 Yest_10, 80-810, 1994
 A; Title: Isolation and DNA sequence of the STE13 gene encoding dipeptidyl aminopeptidase
 A; Reference number: S45451; MUID:95066382; PMID:7973897
 A; Accession: S45451

A; Molecule type: DNA
 A; Residues: 1-931 <ANN>
 A; Cross-references: UNIPARC:UPI0000136060; EMBL:U0000136060; PIDN:g347196; NID:L21944; EMBL:L21944; PIDN:9347196; R; Galisson, F.; Dujon, B.
 Submitted to the EMBL Data Library, October 1995
 A; Description: Sequence and analysis of a 33 kb fragment from the right arm of chromosome XV
 A; Reference number: S60938
 A; Accession: S60946
 A; Molecule type: DNA
 A; Residues: 1-931 <GAL>
 A; Cross-references: UNIPARC:UPI0000136060; EMBL:X92441; PIDN:CAA63182.1; R; Boyer, J.; Fairhead, C.; Gaillou, L.; Galisson, F.; Michaux, A.; Dujon, B.; Thierry, A.; submitted to the Protein Sequence Database, July 1996
 A; Reference number: S67104
 A; Accession: S67112
 A; Molecule type: DNA
 A; Residues: 1-931 <BOY>
 A; Cross-references: UNIPARC:UPI0000136060; EMBL:Z75127; PIDN:g1420507; R; Galisson, F.; Dujon, B.
 A; Experimental source: strain S288C
 A; Title: Sequence and analysis of a 33 kb fragment from the right arm of chromosome XV
 A; Reference number: S71713; MUID:963437977; PMID:8840505
 A; Accession: S71721
 A; Status: nucleic acid sequence not shown; translation not shown
 A; Molecule type: DNA
 A; Cross-references: UNIPARC:UPI0000136060; EMBL:X92441; PIDN:CAA63182.1; R; Galisson, F.; Dujon, B.
 A; Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995
 C; Genetics:
 A; Gene: SGD:STE13; YC11
 A; Cross-references: SGD:S0005745; MIPS:YOR219C
 A; Map position: 15R
 C; Function:
 A; Description: involved in processing of alpha-factor prepropheromone
 C; Keywords: dipeptidylpeptidase; glycoprotein; transmembrane protein
 F:113-150/bound domain: transmembrane #status predicted <TMM>
 F:377/binding site: carbohydrate (Asn) (covalent) #status predicted
 F:785,863,886/Active site: Ser, Asp, His #status predicted
 Query Match Score 726; DB: Length 931;
 Best Local Similarity 26.9%; Pred. No. 4.4e-42;
 Matches 195; Conservative 146; Mismatches 318; Indels 66; Gaps 23;
 Qy 29 FNAYEGNSSLVLENSTDEFGHISINDYSISPDQFPLILEBNVYKQMRHSYTAISYDIVLN 88
 Db 223 FEVNGCNRFLYEGVBPTVSTQIN-YKLDK---LIFGTNLESEFRHSSKGFWIKDLN 277
 Qy 89 K---RQLITEERIPLNNTQ-----WWTWSPVGHKLAYWNNNDIYVKEPLNPMLSRTWT 138
 Db 278 TGNIEPLPPERSDNVYELGSKLSTAHSPNNTYFVYENNTLQQVNSGAKKTTED 337
 Qy 139 GKEDDITVNGITDWTWYREVFSAYSALWSPNGTFLAYAQFNTDVEPLJEYSYSDESLOQV 198
 Db 338 GSXD-IFNAKPKPWYIYEBSVLAQDQIAWPDSSKAVAFRNNTSVDDJRLNNTYNNMEAY 396
 Qy 199 PKTVRVVPKAGAVNPVYKFFUNTDLSSTVNTSQTATASMLGHDYKCDWTWATQ 258
 Db 397 LSDPKIYKPKPSFQNPQFDLFLVN----LQNGIYIYNTSGQK---DSILYNGKWTSP 447
 Qy 259 ERISLQLQURRIONYSYMDICDYDESSERWRNCLVARYHIEMSTGWWGRFR---PBPB 313
 Db 448 DPFRFEITDR-NSKILDVYDIPRSSO-MITVRVNTSNLENGWEYKTDLSIPKPE 503
 Qy 314 FTLDGNSFYKLTISNEEGYRHICYF-QIDKKDCTFTKGTEW---IGIEALTSQDYL 367
 Db 504 LCRMGXGIDIHADSRSQFSLHFLPYTVFAKEPLQTRGNWETVNGIVGEYET-DT1FF 562
 Qy 368 ISNBYKGMGPGRNLNUYKQLSYDTKVTCLSCELNP-ERCQYVSFSKEAKYYQLRCSCGP 426
 Db 563 TANEIGYM---SCHLHSISLTDTNTQFSLQNPDSYDFTPELSSARYITSKLGPD 620
 Qy 427 LPLYTLH-----SSVNDKGJLRLVLEDNSALDKMLQNQVMQPSKXKLDLFLINNE-TKEFWYQ 477

QY	194	ESLQPKTVRVPKAGAVNPTVKFFVVNTDSLSSVTNATSIQTAPPASMLIGDHYLCDV	253	QY	169	NGTFLAYAQFNDETEVPLIEFSYVSDESILQYPKTVRVPYPKAGAVNPTVKFFVVNTDSLSS	228
Db	283	TNFHYPKTF-----AKVLPYI-----TLSTINRKKEPQSRLDQVRDLSHYLLAV	329	Db	216	DESRIVTRVDRGDIVP---RADIGPGGATVNQYPRAGRPNAYVVDLF-----	263
QY	254	TW---ATQBRISLOWLRLRQNTSYMDICDYDESSGRNCLVAROHIE-MMSITGWICRF	307	QY	229	VTNATSIQITAPASMLIGDHYLCDVTW ATQBRISLOWLRLRQNTSYMDICDYDESSGRW	287
Db	330	KWLEINGTEQLOVS-WTNNYONEVALTICD-----WTAICRLEFFYKYASKRWTH-	380	Db	264	VRDLASGRVTAIDLGANKDIIYARVANSADGKTVYVQLSLRDO-KTLDLAAFDATGAG	321
QY	308	RPSBPHFTLDGNSFYKISNEEGYRHICYFQI---DIKQCT-----FI	347	QY	288	NCLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKII-----SNEBGYRHICYFQID	340
Db	381	-----DDFAHSITSPED---TLPFLPHDKDNAFOQQVASLRLSHGQLRTPKFL	425	Db	322	KTILT-----DTDPHFEVSNDFRPLTDGTFLMGSBKGDNQHLYRYAAD	365
QY	348	TKGTMEVIGEALTSDFYLYYISNEYKGMPGRNLYKIQLSDYTKVTCOLSCELN	405	QY	341	KKDCTFTIKGTVWIGEALTSDFYLYYISNEYKGMPGRNLYKIQLSDYTKVTCOLSCELN	400
Db	426	NLGEDVDTSGINGKETRTIFHAAPSKSHRSLSFSYSADENSACISCTI--KNCI	483	Db	366	GKLIAQITKGDWNVIGLEG-----VDAKRYKAIFPSASID	399
QY	406	YYSVSPSKERAKYYQLRCSGPGLPLYTLISSV---NDKGRLRVEDNSALDKMLQNYQMP	460	QY	401	-PERCQYVSVSPEKEAKYQLCRG-----PGELPLYLHSSVNDK	419
Db	484	WAQAOQMDQMKTAIVSCKSPAAPI-----TAIVNLTRMDSDKCTEHANLJYDKYQN-RVB	538	Db	400	TPIERLLEYEVSTAKPGKPKALTISAGGMWAAYVADNGGAFACTYSDPKTPSOTALYSADCK	459
QY	461	SKRUDFIIINETKF-----DKSKYKPLLDDVYASGPCSQKADTVRL	511	QY	440	GIRVLDEOISALD-----KMLQNYQMP-----KKLDFITLNETKFWYQMLPPHFDKSICKY	490
Db	539	EAGLPVIKETIKIISDDFDALIKSIPDIYNRDKHOAPlLVVYGGNDQNTKEATQI	598	Db	460	RYWIEENKLAEQHSPYNPYAAULQPERFGSLKAAD----GET LYTELKPIGFDPAKXY	514
QY	512	NWATYLASTENLIVASFDGSGSGYQGDKLMHAIRNLGFTEVEDEQIAARQFSKM--GTV	569	QY	491	PULLDVYAGPCSOAKDTVFRLNW-----ATYLASTENLIVASFDGRGSQGYQGDKLMHAIN	545
Db	599	GIEEVVASSQAALRIDGEGRSGGRGWYKTRSAVYQGLQTVVEEQIKAIKVVLRLYRL	658	Db	515	PAIVSVYCPHAGRV----MQUWHSPSPRTYEA--GVVIFKLDNRGEGNRSAFKMRALD	568
QY	570	DNKRIAIWGNWSGYYT-SMVLGSGSGFKCGIAVAPUSRWEYVDSVUTERYNGLPTED	628	QY	546	RRUGTFFEVEDQIEARQFSKMGFVFDNKRIAIWGNWSGYGTTSVMSVGSQSGVFKCGIAVAP	605
Db	659	DARRVAFWSXGFFMTLSMVNAPEOFFPKCAYSVAPTNFAYDATERTYNG----DA	714	Db	569	RKLGTVVEVDQJGAKFLASQPVDAKLGTVNGMSYCGFMALMLTAENTPPKAGAACAP	628
QY	629	NLDHYRNSTMRAENFKOVEYLILHGADDNVHFQSAQISKALDVGVDFAMWTFDE	688	QY	606	VSRWEEYDSDVYTERYMLPPTPEDNLDHVRNSTVMSRAENFKOVEYLILHGADDNVHFQQ	665
Db	715	PLESY--SDTYKLDNEKSTRILLMHGLDDNVHFQNSAILDELQRGVDFDLMYFNQ	772	Db	629	PTEWMSLTDIAYTERYMG--KPDENKAGYAYSDINRDKLAPGSSLVNLHGMADDNVFEN	686
QY	689	DHGIASSTAHQIYTHMHSFTIKQCF	713	QY	666	SAQISKALVUDGVYDFQAMWYTDDEDHGJLASSTAH--QHITYHMSH	708
Db	773	AHSLSRSRIS--HVVGKMMTHFLRQCF	795	Db	687	STRIMAAQLQRKATLFEMAMYPGERSHAPSKTGKLSVLUKTHLDPF	731

RESULT 15

AB7516 dipeptidyl peptidase IV [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Accession: AB7516

C:Date: 20-Apr-2001 # sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.B.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwin, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: AB7249; MUID:21173698; PMID:112395647
A:Accession: AB7516
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-738 <STD>
A:Cross-references: UNIPROT:Q9A6B0; UNIPARC:UPI0000007616; GB:AE005673; NID:913423647; E
C:Genetics:
A:Gene: CC2154

Query Match 16.2%; Score 628; DB 2; Length 738;
Best Local Similarity 26.8%; Pred. No. 1.9e-35;
Matches 189; Conservative 115; Mismatches 269; Indels 132; Gaps 23;

QY

52 INDYSISPDQGQFQILLEYN--YUKQWRSYTAQSYDIDLINKRQLITERIPIPNNTQWTWS 108
Db 111 IVEWWDQDQRFILVPLQDGDLYL-----DAVADGKATRLTB-TPGBDVDAVRS 157

QY

109 PVGHKLAYWNNNDIYVKLEPNLPSYRITWTGKBDIYNGITDWYEEVFSAYSALWNSP 168
Db 158 PKGGYVSYTVDQONLYKPKVAGGAETALITDGDALSF-GVAEFIVQEL-DRETYGWNSP 215

Copyright GenCore version 5.1.6
(c) 1993 - 2006 Compugen Ltd.

CM protein - protein search, using sw model

Run on: January 27, 2006, 23:58:33 ; Search time 11 Seconds
(without alignments)
704.874 Million cell updates/sec

Title: US-10-659-055-1_COPY_51_766
Perfect score: 3877
Sequence: 1 NTYRLKLYSLRWTSDHEYLY.....AHQHIVTHMSHFIKQCSLP 716

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 75621 seqs, 10829074 residues

Total number of hits satisfying chosen parameters: 75621

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA_New:
 1: /cgn2_6/podata/2/pubpa/us06_new_pub_pep:
 2: /cgn2_6/podata/2/pubpa/us06_new_pub_pep:
 3: /cgn2_6/podata/2/pubpa/us06_new_pub_pep:
 4: /cgn2_6/podata/2/pubpa/pct_new_pub_pep:
 5: /cgn2_6/podata/2/pubpa/us05_new_pub_pep:
 6: /cgn2_6/podata/2/pubpa/us10_new_pub_pep:
 7: /cgn2_6/podata/2/pubpa/us11_new_pub_pep:
 8: /cgn2_6/podata/2/pubpa/us60_new_pub_pep:
 * Number is the number of results predicted by chance to have a score greater than or equal to the score of the total score distribution, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Match	Length	DB ID	Description
1	3877 100.0	766	6	US-10-522-789-2	Sequence 2, App1
2	3834 98.9	762	7	US-11-116-939-13	Sequence 13, App1
3	2138 55.1	760	7	US-11-186-884-55	Sequence 55, App1
4	5225 13.5	657	7	US-11-151-501-20	Sequence 20, App1
5	272 7.0	657	7	US-11-179-977-1	Sequence 1, App1
6	124 3.2	668	6	US-10-454-377-18	Sequence 118, App1
7	112.5 -3.0	1593	6	US-10-453-372-1092	Sequence 1092, App1
8	112.5 2.9	1436	6	US-10-453-372-1094	Sequence 1094, App1
9	110 2.8	19	7	US-11-116-339-27	Sequence 27, App1
10	110 2.8	72	7	US-11-151-501-22	Sequence 22, App1
11	109.5 2.8	877	7	US-11-077-550-157	Sequence 157, App1
12	109 2.8	2516	6	US-10-647-556A-2	Sequence 2, App1
13	107 2.8	866	7	US-11-077-550-32	Sequence 32, App1
14	107 2.8	871	7	US-11-077-550-155	Sequence 155, App1
15	107 2.8	873	7	US-11-077-550-163	Sequence 163, App1
16	107 2.8	873	7	US-11-077-550-165	Sequence 165, App1
17	107 2.8	876	7	US-11-077-550-66	Sequence 66, App1
18	107 2.8	878	7	US-11-077-550-62	Sequence 62, App1
19	107 2.8	879	7	US-11-077-550-30	Sequence 30, App1
20	107 2.8	908	7	US-11-077-550-64	Sequence 64, App1
21	107 2.8	914	7	US-11-077-550-60	Sequence 60, App1
22	107 2.8	949	7	US-11-077-550-68	Sequence 68, App1
23	107 2.8	1130	7	US-11-077-550-44	Sequence 44, App1
24	107 2.8	1132	7	US-11-077-550-46	Sequence 46, App1
25	106 2.7	878	7	US-11-077-550-12	Sequence 12, App1

ALIGNMENTS

RESULT 1
US-10-522-789-2

Sequence 2, Application US-10522789
; Publication No. US20050260732A1
; GENERAL INFORMATION:
; APPLICANT: TANABE SEIYAKU CO., LTD.
; TITLE OF INVENTION: Three-dimensional structure of dipeptidyl peptidase IV
; FILE REFERENCE: 03-039-PCT
; CURRENT APPLICATION NUMBER: US/10/522-789
; CURRENT FILING DATE: 2005-01-28
; PRIORITY NUMBER: US 60/398,761
; PRIORITY FILING DATE: 2002-07-29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 2
; SEQ ID NO: 2
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-522-789-2

Query Match Score 3877; DB 6; Length 766;
Best Local Similarity 100.0%; Pred. No. 2, 3e-299; Mismatches 0; Indels 0; Gaps 0;

Matches	716; Conservative	Score	3877;	DB 6;	Length	766;
1	NTYKLISLRWISDHEVLYKQENNLILPNAEVGNSSYPLENSTDFEFSHSINDYSISPD 60	1				
51	NTYKLISLRWISDHEVLYKQENNLILPNAEVGNSSYPLENSTDFEFSHSINDYSISPD 110	51				
61	GQFILLEINVKQRHHSATASYDIYDLNKRQLTTERIPNNNTOWTVTSPVGHKLAYWNN 120	61				
111	GQFILLEINVKQRHHSATASYDIYDLNKRQLTTERIPNNNTOWTVTSPVGHKLAYWNN 170	111				

DIVYKIEPLNPLPSITWTKEDITYNGITDWMYEEVEFAYSALWWSPGTFLAYQFND 180

DIVYKIEPLNPLPSITWTKEDITYNGITDWMYEEVEFAYSALWWSPGTFLAYQFND 230

TEVPLIEFSYSDSLSLQPKTKTRPVYPKAGANPTKFKVNTDSLSYYTNATSIQTAP 240

TEVPLIEFSYSDSLSLQPKTKTRPVYPKAGANPTKFKVNTDSLSYYTNATSIQTAP 290

ASMLIGDHYLCDVWATQERISLOWLRRQNLGRNCLVQHQIEMST 300

ASMLIGDHYLCDVWATQERISLOWLRRQNLGRNCLVQHQIEMST 350

TGIVGFRFRESEPHFTLDGNSFYKXISNEEGYRICYFQIDKKDCTFTKGTVETIGIEAL 360

TGIVGFRFRESEPHFTLDGNSFYKXISNEEGYRICYFQIDKKDCTFTKGTVETIGIEAL 410

RESULT 5
US-11-179-977-1
; Sequence 1, Application US/11179977
; Publication No. US20050249789A1
; GENERAL INFORMATION:
; APPLICANT: Genencor International, Inc.
; TITLE OF INVENTION: Alpha/Beta Hydrolase-Fold Enzymes
; FILE REFERENCE: GC511-PCT
; CURRENT APPLICATION NUMBER: US/11/179, 977
; CURRENT FILING DATE: 2005-07-12
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Bacillus
; US-11-179-977-1

Query Match Score 7.0%; Best Local Similarity 20.4%; Pred. No. 6.6e-14; Mismatches 285; Indels 200; Gaps 34;

Matches 153; Conservative 113; MisMatches 285; Indels 200; Gaps 34;

Qy 51 SINDYSISPDQFILLENNYKQPRHSPTRASYDYLNLKQLTEERLPNTTOWV---- 105
Db 15 SVTPQQYAPDGTRAYVKQVNQEKDSITSNIWY-----ETKTGGSVPTWTHGKRR 65

Qy 106 ----TWSPUGHKLAYVMNND----IYKIEPNLPSYRITWKGEDIYNGITIDWWYBEE 156
Db 66 STDPRWSPDORTLAFIGSDREGDAQLYINSTEGBEARLT----DIPY----- 109

Qy 157 VFSAYSLAWMSPNCTFLAYA-----QFNQTEVPLIYFSYFSDSLO----- 197
Db 110 ---GVSKPWLSPDGSILVTISLGESEISDDRE-KTEBDSYEPVEVGLSYKRDGKLT 164

Qy 198 --YPKTRVVPYKAG-----AVNPYTKPFV----VNTDSL----- 227
Db 165 RGAQALVYVS-KEGEMKELTSKHADEGDPAFSPDGKMLVFSANLTEDDAKPHDYI 223

Qy 228 -SVTNATSIQT-----APASMLIGDHYLC-----DVTWATOBRISLQWURRIQNTSV 274
Db 224 MSLESGDQLQTVPHRGSPGSSSFSPDGRYLALLNEKEYKNAT--LSKAWL----- 272

Qy 275 MDICDYYDESSGRVNCNLVAHQHIEINTT----GNYVGRFRSEPEPHTLDGNSFYKTLISNEEG 330
Db 273 -----YDIEGRLCITLTEMVLHDALALIGDSLIGG-AEQRPIWKDQSGFY-VIGTGQ 325

Qy 331 YRHICYFQIDKKDCTFITKGTMWIGIEALTSDYIYKSNEXKGMPGGRNLKYKIQLSNYT 390
Db 326 STGYYISITE--GLVYPIRLEKEYINSLSLSPDQEKFASVTK----- 366

Qy 391 KVTCUSCENPER-CQYVSVSFSKBAKYQLRCRGSGPGLPLYLHSVNDKGRLRVLEDNSA 449
Db 367 -----PDRSELSYSPLGQBEK--QL-----TGANDKPVF-EHTMS 399

Qy 450 LDKMLQNWMQPSKCLDFIILNETKFWQMILPPLPHFDKSKYKDVYAGPCSQKADTVF 509
Db 400 IPBEIQUATE----DGMVNG---W-LMRPAOMEGETTYPLLNIGGPHMYGHYYF 449

Qy 510 RLWRYATLSTENLIVASFDGRGSGYOGCDKIMHAINRRLGTFEVEDQLEAAROFSKNG-F 568
Db 450 H---EFQVLAAGKGYAVVYINPRGSHGQGFBVNAVRGIGGGKDYYDNQAVDAIKRDPH 506

Qy 569 VDNKRRIATNGWSKGCGTSMVLGSGSGPKFCGTAVAPSRWEYDSV----YTERMG 622
Db 507 IDPQRKLGTGSGFMTNWIVGO-TMRKAAVTOQRSTSNWISFHGVSDIGYFFTDNQLE 565

Qy 623 LPTBEDNLHYRNSTVMSAENPKQVEYKLIHGADDVHFQOQAISRALVDGVDEQFA 682
Db 566 HDMFEDTEKLWDRSPLKWAAN--VETPLILHGERDRCPIEQAEQLFTALKMGKETKL 623

Qy 683 MWXTDEDDIGIATASPAHOHITYMHMSHFIQCF 713
Db 624 VRPENASHNL-SRTGHPRQRKIRKLNLYISSWF 653

RESULT 6
US-10-454-437-118
; Sequence 118, Application US/10454437
; Publication No. US20050277115A1
; GENERAL INFORMATION:
; APPLICANT: Pontejus, Markus
; APPLICANT: Kruger, Burkhard
; APPLICANT: Schröder, Hartwig
; APPLICANT: Zeidler, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; FILE REFERENCE: INVOLVEMENT IN HOMEOSTASIS AND ADAPTATION
; CURRENT APPLICATION NUMBER: US/10/454, 437
; CURRENT FILING DATE: 2003-06-13
; PRIORITY APPLICATION NUMBER: US 60/141031
; PRIORITY FILING DATE: 1999-06-25
; PRIORITY APPLICATION NUMBER: DE 19931636.8
; PRIORITY FILING DATE: 1999-07-08
; PRIORITY APPLICATION NUMBER: DE 19932125.6
; PRIORITY FILING DATE: 1999-07-09
; PRIORITY APPLICATION NUMBER: DE 19932126.4
; PRIORITY FILING DATE: 1999-07-09
; PRIORITY APPLICATION NUMBER: DE 19932127.2
; PRIORITY FILING DATE: 1999-07-09
; PRIORITY APPLICATION NUMBER: DE 19932128.0
; PRIORITY FILING DATE: 1999-07-09
; PRIORITY APPLICATION NUMBER: DE 19932129.9
; PRIORITY FILING DATE: 1999-07-19
; PRIORITY APPLICATION NUMBER: DE 19932226.0
; PRIORITY FILING DATE: 1999-07-09
; PRIORITY APPLICATION NUMBER: DE 199322920.6
; PRIORITY FILING DATE: 1999-07-14
; PRIORITY APPLICATION NUMBER: DE 19932292.2
; PRIORITY FILING DATE: 1999-07-14
; REMAINING PRIOR APPLICATION DATA REMOVED - See File Wrapper or P.M.
; NUMBER OF SEQ ID NOS: 442
; SEQ ID NO 118
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-10-454-437-118

Query Match Score 3.2%; Best Local Similarity 19.5%; Pred. No. 0.036%; Mismatches 68; Indels 170; Gaps 23;

Matches 101; Conservative 68; MisMatches 180; DelIns 170; GapS 23;

Qy 297 EMSTGNGCRFRSEPHR-----TLDGNSFYKIIISNEEGYRHYCFOID-KRDCPFT 348
Db 212 DVATHAW---RDSTPGFERTVFSRSLDYFNSETSLTEGG---LYKLVDTCDV 262

Qy 349 KGTWEVIGEALTSDFLYYIS-NEYKGMPGGRNLKYKQLSDVYTKTCLSCELNPERCQY 406
Db 263 KQW-----IFVSPRDTFPIGAP-GLGVLKE----- 290

Qy 407 YSVFSKAKYQYOLRCSPGGLPYLTHSSVNDKGRLRVLEDNSALDKMLQNVOM----- 459
Db 291 ---FLEGGRDF-----PFTPPESTSLGGLATK-NFLVLTENNVTBTIVTPL 337

Qy 460 --PSKKUDFILNE-----TKFNYQMI-----LPPPHDKSKYK 491
Db 338 NDPTEHEDHIDPEHVTAHVWATSPLDDEINWQASAFTEPLLRPLPGALEVKKAP 397

Qy 492 LLIDVYAGPCSKQSKADTVFRLWVATYLAST-----ENIVASFDG----- 530

Qy 398 LQF-----ENAGQETRHOVTSADGCKIPVFTGAFEEFPPVYHAYGGFEVSLTP 449

Qy 531 -----RGSYQGDKIMHAIINRRLGTFEV-EDQIEAARQFSKNG 567

Qy 450 SHSPTRGIAWLEKGGYYFVEANLRGGGEFGPE-WHSQATKLNRMKWEDHRAVLADLVERG 508

Qy 568 FVDNKRKIAIWGNSYGGVTSVNLGSSESGVFKCGIAVAPVSRWEYY----DSVUTTERYMG 622
 Db 509 YATPEQIAIRGSSNGELTSGALTQYPEAFGAIVQVPLADMRLRHTWSAGSMALAEYGN 568
 Qy 623 LPTPEIN--LDHYRNSTVMSRAENPKOVEYLILHGTAADDNYHFEQQSAQSKALDVGVDF 680
 Db 569 PDDPEERAVIEQYSPVQAVVGVEKRIYPALVTTSTRDRVPAHARLFAQALLDAG--- 625
 Qy 681 QAM-WTDEDIGIASSTAHQH----IYTHMHSFIKQ 711
 Db 626 QAVDYYENTECGHAGAIDNKOTAFVESLIVTWIEKTLDQ 664

RESULT 7
 US-10-453-3-72-1092
 Sequence 1092, Application US/10453372
 Publication No. US2006003323A1
 GENERAL INFORMATION:
 APPLICANT: Alsobrook, et al.
 TITLE OF INVENTION: THERAPUTIC POLYPEPTIDES, NUCLEAR ACIDS ENCODING SAME, AND METHOD
 FILE REFERENCE: 21402-589 A
 CURRENT APPLICATION NUMBER: US/10/453,372
 CURRENT FILING DATE: 2001-06-03
 PRIOR APPLICATION NUMBER: 09/789390
 PRIOR FILING DATE: 2001-02-23
 PRIOR APPLICATION NUMBER: 60/185967
 PRIOR FILING DATE: 2000-03-01
 PRIOR APPLICATION NUMBER: 09/833187
 PRIOR FILING DATE: 2001-03-29
 PRIOR APPLICATION NUMBER: 60/195792
 PRIOR FILING DATE: 2000-03-01
 PRIOR APPLICATION NUMBER: 09/839446
 PRIOR FILING DATE: 2001-03-19
 PRIOR APPLICATION NUMBER: 60/139476
 PRIOR FILING DATE: 2000-03-25
 PRIOR APPLICATION NUMBER: 09/863776
 PRIOR FILING DATE: 2001-05-13
 PRIOR APPLICATION NUMBER: 60/208263
 PRIOR FILING DATE: 2000-05-31
 PRIOR APPLICATION NUMBER: 09/939398
 PRIOR FILING DATE: 2001-08-24
 PRIOR APPLICATION NUMBER: 60/227800
 PRIOR FILING DATE: 2000-08-25
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 1609
 SEQ ID NO: 1092
 LENGTH: 1593
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-453-372-1092

Query Match 3.0%; Score 115; DB 6; Length 1593;
 Best Local Similarity 18.1%; Pred. No. 0.61;
 Matches 162; Conservative 118; Mismatches 298; Indels 318; Gaps 43;

Qy 12 WISDEHELYKQENNLLVFNW-----EYGENSSVP-LENSTEDFEGHSIN--DISISPDCQ 62
 Db 124 YTADENYLVLSKGVILHGQKIELENNGRKGISISIDINPELAPSVMMLYSLHPSGE 183
 Qy 63 FIL-LEYYVKQWRHSHSYTASYDIYDLNRQLIFEEIRPNTQWTSPVGHKLAYVNN 120
 Db 184 MVTOSTQFRLEKCPENQNVLNFs-----KEKSTJPGSNIDLQVSRAASNICALWAV 233
 Qy 121 DIIYKIEPN-----LPSYRITWTGKED-----IYNGITDW 152
 Db 234 DQSULLRNYGOLSAQTIVSOLYSPRLGHYFRGLNLBDKLVCPLEDHILNGI--- 289
 Qy 153 YEEEFVSAYSALWNSPNCFLAYAQNDTEVPIE-----YSFSYDESLOYP 199
 Db 290 -----YTP-----AWAFGKDGYDLVTKDPQNRRIFQRQNTTSFRNTQLSFQ 332

Qy 200 KTVRVPYPKAGAVANPVTVKPFVNTDSLSSVNTATSIQ-----ITAPASMLIG-DH 248
 Db 333 --LJSEPMGSDY----WUVKNSRETYTHQAVKRVYLPKFETVDRPQTWTSIDDE 384
 Qy 249 YLCDTWTQERISLQWLRRIONYSVMDIC-DYDESSGRWCN----LYARQHLEMSTTG 302
 Db 385 FQVDFC----AKYNFGQVQGEGTQIRVCREYESSS--NCEKNENBICEQFAOLENG 435
 Qy 303 WVGRFRPSEPHFTLDGNSFY----KISNEGGPHICYFOIDKDCTRIK-GTWEVI 355
 Db 436 CVSQIQTNTKV-PQLRSGLGMFTFHAVIVTESG---TWNQISERTSVPTOLQCT---v 487
 Qy 356 GIBALTSDFDLYYIS-----NEYKGMPGRNLXKIQLD-----YT-----390
 Db 488 NFENDDTPEVRGISYFGTLKFSDPNNPVNPKMNUQLENEDEFIGNYTTDENGAFQFSIDT 547
 Qy 391 ----KVTLSCEBLNPBC-----QYSVFSKEARLT-----QLRC 422
 Db 548 SDIPDPPEFLNKATVTPESCPYLPSWLTPQQLDAH-LVSRFSYTSNSFLKIVPPBKQLEC 606
 Qy 423 SGPDGPLYTHSSVNDKGVLRLVEDNSALDKMQLQNMQPSKKLDLTFINNEKFMWQMLPP 482
 Db 607 NQQ--KVVFYHYSLNE--AYEDDS-----NVKFFFLMMVV-- 637
 Qy 638 -----KGAILLSSCOKERLNKANNGNFSPPISISSLADLAATVYTHPSGEVAD-- 688
 Qy 531 RGSSVQGDK-----INHAINERLGTPEVEDOI 557
 Qy 483 HFDKSKKYPILLD-----VYAGPCSQADTVPRLNWATYLASTENIVASFPG 530
 Db 639 -----RAENPKOVBELLI 653
 Qy 558 EAARQFSKNGFVNWKRIAIWGWSSYGGYTTSMVLSGS-----GVPKCGIAYAPVSPW 609
 Db 748 SAEVSYNNMPSIE-----PGFYFHGLNLDDGKEDPCIPRDMPYNGLYTTPVSNY 798
 Qy 610 EYDYSVITYERYMGLPTPEDNLHDYRNSTM-----RAENPKOVBELLI 653
 Db 799 GDGJINVTVRNMGKV-FTNL-HYTRKPEKIMVQCVVFRLEHVSGIRENADYEQALL 856
 Qy 654 HGTTADD-----NVEHQQSQIISKALVDPVDFQAMWY---DEDRGIASST 696
 Db 857 QTVRTNFPETWMMDLVSVDSGSANLSFLIPDTITQWEASGFCCVNGDVGFISSTT 912

RESULT 8
 US-10-453-372-1094
 Sequence 1094, Application US/10453372
 Publication No. US2006003323A1
 GENERAL INFORMATION:
 APPLICANT: Alsobrook, et al.
 TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHO
 FILE REFERENCE: 21402-589 A
 CURRENT APPLICATION NUMBER: US/10/453,372
 CURRENT FILING DATE: 2003-06-03
 PRIOR APPLICATION NUMBER: 09/789390
 PRIOR FILING DATE: 2001-02-23
 PRIOR APPLICATION NUMBER: 60/185967
 PRIOR FILING DATE: 2000-03-19
 PRIOR APPLICATION NUMBER: 60/139476
 PRIOR FILING DATE: 2000-03-25
 PRIOR APPLICATION NUMBER: 09/863776
 PRIOR FILING DATE: 2001-05-13
 PRIOR APPLICATION NUMBER: 60/208263
 PRIOR FILING DATE: 2000-05-31
 PRIOR APPLICATION NUMBER: 09/939398
 PRIOR FILING DATE: 2001-08-24
 PRIOR APPLICATION NUMBER: 60/227800
 PRIOR FILING DATE: 2000-08-25
 Remaining Prior Application data removed - See File Wrapper or PALM.

```

; PRIOR APPLICATION NUMBER: 60/227800
; PRIOR FILING DATE: 2000-08-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1609
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO: 1094
; LENGTH: 1436
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-452-372-1094

Db 752 SSNT 755

RESULT 9
US-11-116-939-27
; Sequence 27, Application US/11116939
; Publication No. US20050265995A1
; GENERAL INFORMATION:
; APPLICANT: Stephen Tomlinson
; APPLICANT: Richard J. Quigg
; TITLE OF INVENTION: TISSUE TARGETED COMPLEMENT MODULATORS
; FILE REFERENCE: 19113.0115U2
; CURRENT APPLICATION NUMBER: US/11/116,939
; CURRENT FILING DATE: 2005-04-28
; PRIORITY NUMBER: 60/565,907
; PRIORITY FILING DATE: 2004-04-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSEQ For Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; note=synthetic
; OTHER INFORMATION: Construct
US-11-116-939-27

Query Match 2.9%; Score 112.5; DB 6; Length 1436;
Best Local Similarity 18.1%; Pred. No. 0.83; Indels 309; Gaps 40;
Matches 153; Conservative 109; Mismatches 273;

Qy 55 YSISPDGQFL--LYNNYRQWRRHSTASYDIYDLNRKOLITEERIPPIINTQWYWSPVGH 112
Db 19 YSLHPGGEMVTDSIOPRIERCFENOVNLNES-----KEKSLPGNIDLOVSAASN 68
Qy 113 KLAYTWNNDIYVKLEPN-----LPSRITWTGKED---II 144
Db 69 SLCALAWAVDGSVLLRNIGLSAQTVISQLYSREHHGYYFRGLNLEDGLKVPCLDEIL 128
Qy 145 YNGITDWWYBEEVEFAYSALWWSPTGTFLAYAQOFNDEPYLIE-----YSFY 191
Db 129 YNGI-----YYP-----AWADRGKDGYDLVVDPQNRRIFQRONVTSFR 167
Qy 192 SDESLOYPKTIVRVPYKAGAVNPVYKFVVNTDSLSSVTNATSQ-----ITAPA 241
Db 168 NITQLSFO----LISPMPFGDY----WIVVYKRNSRETWTQFAVKRYVLPKFEVTVNAPQ 219
Qy 242 SMLIG-DHYLCDVTAQTERISLOWLRIRONYSMIDC-DYDESSGRWC-----LVARQ 294
Db 220 TVTSDDEFQDVVC----AKYNGQPVQGETQIRVRCYEFSSS---NCEKNEEIEBQ 270
Qy 295 HIEMSTTGMGRFRPSEPHFTLDGNSFY----KLIISREGYRHICYFQIDKEDCTFTIK 349
Db 271 FIAQLENQCVSQINTKV-FQLYRSGLFLMTFHRAVIVTESSG---TVMQISEXTSVFTQ 325
Qy 350 --GIVEVIGBALTDLYIS-----NEYKGMPGGRANLYKIQSLD-----YT---- 390
Db 326 LLGT---VNPFNMDDTFYRGEISYFGLKESDPNNVPMNKLLQLENEFIGNYTTDENG 382
Qy 391 -----KTVCLSCELNPERC-----QYYSVSFSKEAKYV----- 418
Db 383 EAQFSTDSDIFDPBNKLATYVRFESTCPLPSWLTPOQLDAHF-LVSRPFSRSTSFLXIV 441
Qy 419 ---QLRCSPGPGLPLYLTHSSVNDKGRLVLEEDNSALDKMLQNVQMPSKLLDFILNETKF 474
Db 442 PEPKOLECNQQ---KVTWVHYSLNSE---AYEDDS-----NVKF 474
Qy 475 WYQMLLPHPDKSKYPLLD-----VIAGPQSKQADTVFLRINWATLASTEN 522
Db 475 FYLMNV-----KGAIILSGQKBIRNIKAANGNFSPFISADLAAPAVLFYTLHPSG 526
Qy 523 IIIVASFDGRGSSGYQCDK-----IMHAINRIG 549
Db 527 EIVAD---SVRFQYDKFCRHKVNKFKSNEQGLPICALRAVRNL 582
Qy 550 TFEYBDQIEARQFSKMGFVDNKRKIAWGRSYGGVTSNVLGSGS-----GVFKCGI 601
Db 583 LLKSSQQLSAESVNNVNPSTB-----PYGFYFHGNLDDGKEDPCIPQRDMFNGL 633
Qy 602 AVAPYSRWEYDSSVTERYMLPPIPDDHYRSTVMS-----RAENP 645
Db 634 YYTPVSNYGDGDIYNNVTRANGLKV-FTNL-HYRKPEKLMVQCVYFLHLVASSGIRGENA 691
Qy 646 KQVEYLLIHTADD-----NVHFOOQAQSALKAYDVGVDQAMWYT---DEDIGI 692
Db 692 DYQEQAIIQTVRTRNFPETNWMDLVSVDSSGSANLSPFLPDTIOWEASGGFCVNGDVGFGI 751
Qy 693 ASST 696
Db 752 SSNT 755

```

Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 45
 SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO: 22
 LENGTH: 72
 TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:
 OTHER INFORMATION: consensus sequence

US-11-151-601-22
 Query Match 2.8%; Score 110; DB 7; Length 72;
 Best Local Similarity 48.3%; Pred. No. 0.022; Gaps 0;
 Matches 29; Conservative 3; Mismatches 28; Indels 0;

Qy 524 IVASFGRGSGCYGDKIMHAINRRLGTFEVDQEAAARQFSKMGEVNDNCRIAIAGWSYGG 583

Db 11 IYAVVDIRGGGEYGYOKRWHEATRLLKCNFNDTAAAEYLSQLGTTSPKRIAIFCGSNGG 70

RESULT 11
 US-11-077-550-157

Sequence 157, Application US-11077550
 Publication No. US20050244435A1

GENERAL INFORMATION:

APPLICANT: Shone, Clifford Charles

APPLICANT: Quinn, Conrad Padraig

APPLICANT: Foster, Keith Alan

APPLICANT: Chaddock, John

APPLICANT: Marks, Philip

APPLICANT: Sutton, J. Mark

APPLICANT: Stanccombe, Patrick

APPLICANT: Wayne, Jonathan

TITLE OF INVENTION: Recombinant Toxin Fragments

FILE REFERENCE: 1581-0130004

CURRENT APPLICATION NUMBER: US-11/077,550

PRIOR APPLICATION NUMBER: 10/241,596

PRIOR FILING DATE: 2002-09-12

PRIOR APPLICATION NUMBER: 09/255,829

PRIOR FILING DATE: 1998-02-13

PRIOR APPLICATION NUMBER: PCT/GB97/02273

PRIOR FILING DATE: 1997-08-22

PRIOR APPLICATION NUMBER: 01/782,893

PRIOR FILING DATE: 1996-12-27

PRIOR APPLICATION NUMBER: GB96759996.5

PRIOR FILING DATE: 1996-12-13

PRIOR APPLICATION NUMBER: GB9617671.4

PRIOR FILING DATE: 1996-08-23

NUMBER OF SEQ ID NOS: 179

SOFTWARE: PatentIn version 3.1

SEQ ID NO: 157

LENGTH: 877

TYPE: PRT

ORGANISM: Clostridium botulinum

US-11-077-550-157

Query Match 2.8%; Score 109.5; DB 7; Length 877;

Best Local Similarity 19.5%; Pred. No. 0.73; Gaps 32;

Matches 144; Conservative 91; Mismatches 259; Indels 243; Gaps 32;

Qy 30 NAEYGNSSV--FLENSTPDEFGHISINDYSISPDQGQFILEINNVKQNRHSYTASYD 86

Db 250 NAYENMSGLEVSFEELRTFG--GHDAKF1DSLQNEFRYYNNFKF-----DAS 297

Qy 87 -LNKQLQILITERIPNTQWNTWSPVGHCLAYWNNDIVTYIEPNLPSRYITWTKEDDIY 145

Db 298 TLNKAKSIV---GTTASLOQMNPKVKEYLSEDTSCKFSVDKLF-----DCLY 344

Qy 146 NGITDWWYEEBFVFSAYSALWWSPGNTFLAY--AGFNDFEVPLIEYSFSDESLQYPKTVR 203

Db 345 KMLTB-ITYEDNFVFKFV--LNKRTYLNFDKADEVKINIVPKNVITYDGFNLRTNU-- 399

	Qy 204 VPYKAGAVNPTVKEPVVNTDSLSVTNATSIQXTAPASMLIGDHY--LCDVTWATOERI 261	Qy	
Db	400 -----ANNENGQTTEINNMNFYKLKNPTG-----LEFYKLUVGRLITSKTK 442	Db	
Qy	262 SL-OWLRRIONYSWMDICDYDESSGRWMNCLVAROHIEMTSTGWWGRFRSEPHFTLDGNS 320	Qy	
Db	443 SLDGCKYK1EGRANLDC-----IKVNNDLFLFSEDFNTNDLNK 483	Db	
Qy	321 FYKLISN--EEGYRHICYFQIDKDCTFTKGTEVIGEALTSDYL--YYISNEVKGM 375	Qy	
Db	484 GEEITSDTNTEAAENINSLDIOQYLTLTNFDNPENISIENLSDDIGOLELMPNIERF 543	Db	
Qy	376 PGGRMLYK1QLSDYTKTVCCLSCEANPERQYSISFSKSAKYQLRCRGPGPLPYTLHSS 435	Qy	
Db	544 PNGK--KVELDKYTMEMHL-----RAQEPEFGKSRTA-----LTON 577	Db	
Qy	436 VNDGLRLVEDNSALDKMLQNQVNPMSKDFLTLNETKEWYQMLLPHPFDKSYPLLLD 495	Qy	
Db	578 VNEALLN-----PSRVYTFP-----592	Db	
Qy	496 VYAGFCOSOKADTPVRNAT-----YLASTENITIVASPDPGRGSYOGDKTMH-----542	Qy	
Db	593 -----SSDVYKKVNATEAAMFLGWVQLVYDFDTDTSESVSTDKLAIDIILIPYIG 644	Db	
Qy	543 -AINRRLGTFEEVEOIEARQFS---KMGFDVNDKRIAIWG-WSYGGYUTSTMVNGSGSGV 596	Qy	
Db	645 PALN----IGNMLYKQDFVGALIFSGAVILLEFIPBIAIPVGLTPALVSYANKVUTVQT--700	Db	
Qy	597 FKCGIAVAPVSRWYDVSYTTERMGLPTPEDNLDHYNSTMRAENFKQVEYLHGT 656	Qy	
Db	701 ---IDNALSKRNKEWDEVY-----KVITVNWL 724	Db	
Qy	657 ADDNVH-----FOQSQAQTSKALYDVGVDEOAMVYDDEDHG-----IASSTAH 698	Qy	
Db	725 AKVNTQIDLJRKMKREALENQAEATKAI---INYQNOYTEBKRNNINFNDLSSKLN 780	Db	
Qy	699 QHI--YTHMSHF1KQC 712	Qy	
Db	781 ESTINKAMNINKFNQ 797	Db	
RESULT 12			
	US-10-647-956A-2		
	Sequence 2, Application US-10647956A		
	; Publication No. US20050251878A1		
	GENERAL INFORMATION:		
	; APPLICANT: ffirach-Constant, Richard		
	; APPLICANT: Bowen, David		
	; APPLICANT: Rocheleau, Thomas		
	; APPLICANT: Waterfield, Nicholas		
	; TITLE OF INVENTION: DNA SEQUENCES FROM PHOTORHABDUS LUMINESCENS		
	; FILE REFERENCE: 61645		
	; CURRENT APPLICATION NUMBER: US-10/647,956A		
	; CURRENT FILING DATE: 2003-08-26		
	; PRIOR APPLICATION NUMBER: CURRENT APPLICATION NUMBER: US/09/817,514		
	; PRIOR FILING DATE: CURRENT FILING DATE: 2000-03-26		
	; PRIOR APPLICATION NUMBER: US 60/191806		
	; PRIOR FILING DATE: 2000-03-24		
	; NUMBER OF SEQ ID NOS: 8		
	; SOFTWARE: PatentIn version 3.0		
	; SEQ ID NO: 2		
	; LENGTH: 2516		
	; TYPE: PRT		
	; ORGANISM: Photorhabdus luminescens		
	US-10-647-956A-2		
Query	20 YKQENNLY--FNAE-----YGNSS-----VPLENSTPDEFGHISINDYSISPDQGQFILEINNVKQNRHSYTASYD 86	Query	2.8%; Score 109; DB 6; Length 2516;
Db	250 NAYENMSGLEVSFEELRTFG--GHDAKF1DSLQNEFRYYNNFKF-----DAS 297	Db	Best Local Similarity 18.5%; Pred. No. 3.4%; Mismatches 205; Indels 310; Gaps 39;
Query	87 -LNKQLQILITERIPNTQWNTWSPVGHCLAYWNNDIVTYIEPNLPSRYITWTKEDDIY 145	Query	20 YKQENNLY--FNAE-----YGNSS-----VPLENSTPDEFGHISINDYSISPDQGQFILEINNVKQNRHSYTASYD 86
Db	298 TLNKAKSIV---GTTASLOQMNPKVKEYLSEDTSCKFSVDKLF-----DCLY 344	Db	Matches 137; Conservative 89; Mismatches 205; Indels 310; Gaps 39;
Query	146 NGITDWWYEEBFVFSAYSALWWSPGNTFLAY--AGFNDFEVPLIEYSFSDESLQYPKTVR 203	Query	20 YKQENNLY--FNAE-----YGNSS-----VPLENSTPDEFGHISINDYSISPDQGQFILEINNVKQNRHSYTASYD 86
Db	345 KMLTB-ITYEDNFVFKFV--LNKRTYLNFDKADEVKINIVPKNVITYDGFNLRTNU-- 399	Db	Matches 137; Conservative 89; Mismatches 205; Indels 310; Gaps 39;
	1239 YQGEDTLLVMFYRNQDFTLDSYKRNAMSMQGLYIFADMASKDMTPEGSVYR--DMSYQQFD 1295		

QY 68 YNYVKQRHSYTASYDI-YDLNRKQ-----LITERIPINTQWTVTWSPGHKLAY 116
Db 1296 TNNVRVNNRYAEDVEIPISSRSRKYGMDYIISMVNGDIFT-----INYKA- 1345
QY 117 VNNNDIVKKEPNLPSYRTWTGREDIINGITDWYE-----EVRSA 160
Db 1346 --SSDLKIYVSPKL-----RIHNG-----YEGRNRCNLMKYGLGDKTV 1387
QY 161 YSALWNSPNT--FLAYAQFNDETEPVLIBYS-----FYSDESOYLPKTVR- 203
Db 1388 YTSGLYNPNNSNKLMPY-----PVQSGNTSGLNQGRLLFHRTD--YPSKVZAW 1437
QY 204 VPYPKAGAVNPT-----VKEFVUNTDSSUVTNATS-IQT---APAS 242
Db 1438 IGAKRSLTNQNAALTGDATDSINKPDDLKQYIFMTDSKGTAATDVSGVEINATAISAK 1497
QY 243 MLI----GDHYLCDTWATERISIQ-----WLRRIQNTSYMDI- 277
Db 1498 VQIIVKAGGK--EQTFADKDVSIOPSPSFDEMKNYQFNALIDGSGLNFINNSASIDYT 1554
QY 278 -----CDYDESS-----GRNCLVARCHI 236
Db 1555 FTAFALDKRLGYYESFSIPTLKVSTDNALTHHNENGQAQYOMQOSYRTRLNTLFARQLV 1614
QY 297 EMSTIG-----WVGRFRPSEPHFTLDGNSFY-----KILSN 327
Db 1615 ARATGIDTILSMETQNIOEPQL--GKGFYATFVPPNLSTHGDERMFKLTKHVVN 1671
QY 328 EGGYRHICY-----FOID-----KKDCTF--ITKGTW-----E 353
Db 1672 NS--HIVSGQLDTNNINTLFLPDYPLNQDYHAKVYMTFKSPSGTGWGPHEYRD 1728
QY 354 VIGEAL-----TSDYLYISNEYXKGMPGCRNLKYIQSLQSDTYKTCSELNPER 403
Db 1729 DKGIVTINPKSILTHPEVNLNISSEPMDFSGANSLYFWELFFYTPMLVAQRLLHQ- 1787
QY 404 CQYVSFSFSKEAKCYQLRCSPGGLPYLTLSHSSYNDKGRLV--LEDNS-----448
Db 1788 -----NFDEANRWKYVNSPSG--YIVHGQIONYQNNVRPLIEDTSNSPDLDSVPDD 1838
QY 449 ---ALDKMLQNVQMPSKLDFII-----INETKFWYQMLLPPHFDKSCKYP 491
Db 1839 AVAQHDPMHYKVSTEMRMLDILLIARGDHAYROLERDTNEAKWYMOAL--HHLGDKYL 1896
QY 492 LLLDVYAGPESQKADTFVRLN 512
Db 1897 PLSTTWSDPDLRADITION 1917

RESULT 13
US-11-077-550-32
; Sequence 32, Application US/11077550
; Publication No. US2005024435A1
; GENERAL INFORMATION:
; APPLICANT: Shore, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, Philip
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stancome, Patrick
; APPLICANT: Wayne, Jonathan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; CURRENT APPLICATION NUMBER: US/11/077,550
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: 10/241,596
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 09/555,829
; PRIOR APPLICATION NUMBER: PCT/GB97/02273
; PRIOR FILING DATE: 1997-08-22

; PRIORITY: PatentIn version 3.1
; SEQ ID NO: 32
; LENGTH: 866
; TYPE: PRT
; ORGANISM: Clostridium botulinum
; US-11-077-550-32
Query Match 2.8%; Score 107; DB 7; Length 866;
Best Local Similarity 19.6%; Pred. No. 1.1; Mismatches 261; Indels 242; Gaps 31;
Matches 144; Conservative 89; Score 107; DB 7; Length 866;
Qy 30 NAEYGNSSV--FLENSTDFEHSINDYDISISPDQFILLEMYVQMRHSYTASYDIYD 86
Db 248 NAVYEMSGLEVSPEELRPG--GHDAKFDPSLQNEFRLYYNKF-----DIA 295
Qy 87 -LNKROLITEERIPNNTQWTVTWSPGHKLAYYWNNDLVVKIEPNLPSYRITWTGKEDITY 145
Db 296 TLRKAKSV-----GTTASLQTMKVNPKBYKYLILEDSTGKFSDYKFLKE-----DLY 342
Qy 146 NGITDWTEEVEEVSAYSAIWWNSNGNTFLAY--AQNFNTEVPLIEFSYSDESLOQPKTR 203
Db 343 KMLTE-IXTEDNPKFKPV--LNRKTYLNFDFKAVFKINVPKWNTYTDGFNLRTNL-- 397
Qy 204 VPKPKAGAVNPYTFKFVNTDSLUSVTNATSIOTAPASMLGHDY--LCDVTVWATQRI 261
Db 398 -----ANAFNQONTENNNPNTKLXNFTG-----LFEFYKLQVRCIUTTSK 440
Qy 262 SLOWLRRQNTSYMDICDYEDESSGRWNCLVARQHIEMTTGMYGRFRPSEPHFTLDGNSF 321
Db 441 SLDDDGYNKALNDLC-----IKVNNDLFFPSEDNTNDLNG 481
Qy 322 YKISN--EGGRHICYPQIDKDCDTRITKGTVWEGIGEALTSIDL--YYISNEYKGP 376
Db 482 EELTSDTNEAABENISSLDIOQYLTENFNEPENTSIENSDDIGOLEMPNIERFP 541
Qy 377 GGRNLVYKIQSLQSDTYKTCVTLSCELNPERQOYSSFSKEAKYVQLCSPGGLPLYTLHSV 436
Db 542 NGK--KYEKDLYTMFL-----PAQEFPKSRIA-----LNTNSV 575
Qy 437 NDKGRLVLEDNSALDKMLQNVQMPSKLDFITNETKFWYQMLLPPHFDKSCKYPLLUDV 496
Db 576 NEALLN-----PSRVTFP-----589
Qy 497 YACFCSQKADTVPLNAT-----YLASTENIIVASDFGRGSGYQGDKLMH-----542
Db 590 -----SSDYVKVNKATEAAAMFLGWYDQYDFTDETSEVSTDKIADITIIPYGP 642
Qy 543 AINRRLGTFEVEDQIEAROF-----KMGFDVNDRKRIAIGW-WSYGGYTTSMSVMLGSGSGVF 597
Db 643 RN--IGMMLYKODFVGALIFSGAVILLEFIPRIAPVLTGPAVLSVANKVLTQV---697
Qy 598 KCGIAVAPVSRWEYDSVYSTERIMGLPTPDEDNLHYRNSTMRSRAENFKQVEYLHGT 657
Db 698 ---IDNALSKRNEKWDEVY-----KVITVNL 722
Qy 658 DDWVH-----FQQSQTISKALVDVGVDFOQAMWYTDDEHG-----IASSTAHO 699
Db 723 KVNTQDILJRKQKREALENQAEATAK-----INYQNOYTEEBSKNNINFIDLOSSKLINE 778
Qy 700 HI--YTMMSHFYKQC 712
Db 779 SINKAMNINKEFNC 794

Sequence 155, Application US/11077550
 Publication No. US2005024435A1
 GENERAL INFORMATION:
 i APPLICANT: Shone, Clifford Charles
 i Quinn, Conrad Padraig
 i Foster, Keith Alan
 i Chaddock, John
 i APPLICANT: Marks, Philip
 i APPLICANT: Sutton, J. Mark
 i APPLICANT: Stancome, Patrick Wayne, Jonathan
 i TITLE OF INVENTION: Recombinant Toxin Fragments
 i FILE REFERENCE: 1581 0130004
 i CURRENT APPLICATION NUMBER: US/11/077, 550
 i CURRENT FILING DATE: 2005-03-11
 i PRIOR APPLICATION NUMBER: 10/241, 596
 i PRIOR FILING DATE: 2002-09-12
 i PRIOR APPLICATION NUMBER: 09/255, 829
 i PRIOR APPLICATION NUMBER: PCT/GB97/02273
 i PRIOR FILING DATE: 1997-08-22
 i PRIOR APPLICATION NUMBER: 08/782, 893
 i PRIOR FILING DATE: 1996-12-27
 i PRIOR APPLICATION NUMBER: GB965996.5
 i PRIOR FILING DATE: 1996-12-13
 i PRIOR APPLICATION NUMBER: GB9617671.4
 i PRIOR FILING DATE: 1996-08-23
 i NUMBER OF SEQ ID NOS: 179
 i SOFTWARE: PatentIn version 3.1
 i SEQ ID NO: 155
 i LENGTH: 871
 i TYPE: PRT
 i ORGANISM: Clostridium botulinum
 US-11-077-550-155

Query Match 2.8%; Score 107; DB 7; Length 871;
 Best Local Similarity 19.3%; Pred. No. 1.1;
 Matches 147; Conservative 94; Mismatches 263; Indels 258; Gaps 34;

Qy 16 HEYLKQE-----NNLVNAEYGNNSYLENS---TFDEP---GHSINDYSISPD 60
 Db 223 HELIYAGRHLIGIAINPNREVKV--NNTAYNENGLSEEVRLTGGHDAKEPTIDSQ 279
 Qy 61 GQFILENYTKWHRHTASYDIY-LNKROLITEER1PNNTQWNTSPVGHKLAYTN 119
 Db 280 NEFRILYYNFKF-----DIASTLINKAKSIV----GTAASLQNMKNVPEKYL 324
 Qy 120 NDIVVKEPNLPSYRITWKGEDIYNGTIDWVYEEVSAYSALWWSPNGTFLAY--AQ 177
 Db 325 EDTSGKFSVDKLFK-----DLYKMLTE-IYEDNTVFKFKV--LNRKTYLNFDKAV 373
 Qy 178 FNDTEVPLIEYSFSDESLOYPKTPVRYPKAGAVNPVTFKFFVNTDDSLSSVTNATSQI 237
 Db 374 FKINTVPKNTYIYDGFDNIRNTN-----AANFNQQTNEINNMFTKLKNFTG---421
 Qy 238 TAPASMLGDHY--LCDWTWATQRISLQMLRRQNYSYMDCYDDESSGRNCLVARQH 295
 Db 422 -----LFEFYKLLCVRGJLITSKTSLD---KSYKNKALNDLC-----454
 Qy 296 IEMSTTGWGRFRSEPHFTLDGNSFYKLSN---EESGRHICYFOIDRKDCPFITKGTW 352
 Db 455 -IKVNNDLFFSEDFNDLNKEETSDINAEABENISLDLQQYLTTFNFDNEP 512
 Qy 353 EVIGEATLSDYL--YYISNEYKMPGCGANLYKLQSYTKVTCNLSCENPNCQYSV 410
 Db 513 ENISTENSLSDIOLGLBMPNIEFPNGK---KVBLDKTMFLY-----RAQEFPHG 562
 Qy 411 FSKEAKYQQLRCRSGPGLPLYTLHSSVNDKGLRVEDNSALDKMLQNYQMPSKLDPILN 470
 Db 563 KSRIA-----LTVSNVALLN-----PSRVYTF---586
 Qy 471 ETKFWYQMIILPHEFDKSCKYPLLDVYAGPCSKQADTVYERLNAT---YLASTENTIVA 526
 Db 587 -----SSDVTKVKVNKATEAMPLGWVTEQLVVD 613
 Qy 527 SFDRGSGYQGDKIMH-----A1NRRLGTEVEOQEAEARQFS---KMGFVDNK 572
 Db 614 FTDSETSEVSTTDKAIDITLIPYGPALN--IGNMLYKODFVGALIFSGAVILLEPIEI 671
 Qy 573 RIAIWG-WSYGYYTSMVNLGSGCFVKCGJAVAPSRWEYDSUYTERMGLPPPEDNLD 631
 Db 672 AIPVLTGFALVSYANKVLTVO-----IDNALSKRNKWDFTY-----710
 Qy 632 HYRNSTMMSRAENPKQVEYLILLSTADDNVH-----FOOSAQISKALVDVGVDP 680
 Db 711 -----KYIVTNWLAKTNTQDILRKOMKEALENQAETKAI---INY 749
 Qy 681 QAMWYTDDEHQ-----IASSTAHOHI---YTHMSHF1KOC 712
 Db 750 QYNQYTEEKKNNINFDDLSKINESINKMINFKLNQC 791
 RESULT 15
 US-11-077-550-163
 Sequence 163, Application US/11077550
 i Publication No. US2005024435A1
 GENERAL INFORMATION:
 i APPLICANT: Shore, Clifford Charles
 i Quinn, Conrad Padraig
 i APPLICANT: Foster, Keith Alan
 i APPLICANT: Chaddock, John
 i APPLICANT: Marks, Philip
 i APPLICANT: Sutton, J. Mark
 i APPLICANT: Steancombe, Patrick
 i APPLICANT: Wayne, Jonathan
 i TITLE OF INVENTION: Recombinant Toxin Fragments
 i FILE REFERENCE: 1581.0130004
 i CURRENT APPLICATION NUMBER: US/11/077, 550
 i PRIOR APPLICATION NUMBER: 10/241, 596
 i PRIOR FILING DATE: 2002-09-12
 i PRIOR APPLICATION NUMBER: 09/255, 829
 i PRIOR FILING DATE: 1999-02-23
 i PRIOR APPLICATION NUMBER: GB97/02273
 i PRIOR FILING DATE: 1997-08-22
 i PRIOR APPLICATION NUMBER: 08/782, 893
 i PRIOR FILING DATE: 1996-12-27
 i PRIOR APPLICATION NUMBER: GB9625996.5
 i PRIOR FILING DATE: 1996-12-13
 i PRIOR APPLICATION NUMBER: GB9617671.4
 i PRIOR FILING DATE: 1996-08-23
 i NUMBER OF SEQ ID NOS: 179
 i SOFTWARE: PatentIn version 3.1
 i SEQ ID NO: 163
 i LENGTH: 873
 i TYPE: PRT
 i ORGANISM: Clostridium botulinum
 US-11-077-550-163

Query Match 2.8%; Score 107; DB 7; Length 873;
 Best Local Similarity 19.3%; Pred. No. 1.1;
 Matches 147; Conservative 94; Mismatches 263; Indels 258; Gaps 34;

Qy 16 HEYLKQE-----NNLVFNAYGNSVLENS---TFDEP---GHSINDYSISPD 60
 Db 225 HELIYAGRHLIGIAINPNREVKV--NNTAYNENGLSEEVRLTGGHDAKEPTIDSQ 281
 Qy 61 GQFILENYTKWHRHTASYDIY-LNKROLITEER1PNNTQWNTSPVGHKLAYTN 119
 Db 286 GQFLLTENYKVKRMRHSYTASYDIY-LNKROLITEER1PNNTQWNTSPVGHKLAYTN 119
 Qy 282 NEFRILYYNFKF-----DIASTLNKAASSV-----GTTAISLQMYKNTFKYLLS 326
 Db 327 EDTSGKFSVDKLFK-----DLYKMLTE-IYEDNTVFKFKV--LNRKTYLNFDKAV 375
 Qy 378 ENDTEVPLIEYSFSDESLOYPKTPVRYPKAGAVNPVTFKFFVNTDDSLSSVTNATSQI 237
 Db 421 FKTINTVPKNTYIYDGFDNIRNTN-----AANFNQQTNEINNMFTKLKNFTG---586
 Qy 478 TAPASMLGDHY--LCDWTWATQRISLQMLRRQNYSYMDCYDDESSGRNCLVARQH 295
 Db 422 -----LFEFYKLLCVRGJLITSKTSLD---KSYKNKALNDLC-----454
 Qy 296 IEMSTTGWGRFRSEPHFTLDGNSFYKLSN---EESGRHICYFOIDRKDCPFITKGTW 352
 Db 455 -IKVNNDLFFSEDFNDLNKEETSDINAEABENISLDLQQYLTTFNFDNEP 512
 Qy 353 EVIGEATLSDYL--YYISNEYKMPGCGANLYKLQSYTKVTCNLSCENPNCQYSV 410
 Db 513 ENISTENSLSDIOLGLBMPNIEFPNGK---KVBLDKTMFLY-----RAQEFPHG 562
 Qy 411 FSKEAKYQQLRCRSGPGLPLYTLHSSVNDKGLRVEDNSALDKMLQNYQMPSKLDPILN 470
 Db 563 KSRIA-----LTVSNVALLN-----PSRVYTF---586
 Qy 471 ETKFWYQMIILPHEFDKSCKYPLLDVYAGPCSKQADTVYERLNAT---YLASTENTIVA 526
 Qy 587 -----SSDVTKVKVNKATEAMPLGWVTEQLVVD 613

Search completed: January 28, 2006, 00:05:16

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2006, 14:06:40 ; Search time 163 Seconds
(without alignments)

Title: US-10-659-055-1_COPY_51_766
Perfect score: 3877

Sequence: 1 NTYRLKLYSLRWSIDHBYL AHQHITHMMSFIKQCFSLP 716

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0*

Maximum Match 100*

Listing first 45 summaries

Database : UniProt 05.80:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3877	100.0	766	1 DPP4_HUMAN	P27487 h dipeptidyl peptidase 4 (EC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP4) Q53TNA_HUMAN
2	3877	100.0	766	2 Q5R7G7_PONY	Q5R7G7 pony pygmaea
3	3848	99.3	765	1 DPP4_BOVIN	P81425 b dipeptidyl peptidase 4 (EC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP4) Q9N17_f_dipeptidyl peptidase 4 (EC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP4) Q9N17_f_dipeptidyl peptidase 4 (EC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP4)
4	3531	91.1	765	1 DPP4_FELCA	P24411 s dipeptidyl peptidase 4 (EC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP4) Q9N17_f_dipeptidyl peptidase 4 (EC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP4)
5	3501	90.3	765	1 DPP4_PIG	P14140 r dipeptidyl peptidase 4 (EC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP4) Q9N17_m_dipeptidyl peptidase 4 (EC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP4)
6	3489	90.0	766	1 DPP4_RAT	P28843 m dipeptidyl peptidase 4 (EC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP4) Q5z181_gallus_gallus
7	3345	86.5	767	1 DPP4_MOUSE	Q75583 agkistrodon piscivorus
8	3340	86.1	760	1 DPP4_XENIA	P70092 xenopus laevis
9	2490	64.2	751	2 Q75582_AGRIB	Q75582 agkistrodon piscivorus
10	2489	64.2	751	2 Q52181_CHICK	P97321 mus musculus
11	2478	63.9	759	2 Q641D5_XENIA	Q91651 xenopus laevis
12	2309	59.6	737	2 P70092_XENIA	Q4rcj4 tetrodon natus
13	2292	59.1	748	2 P70092_XENIA	Q53TP5_HUMAN
14	2141	55.3	760	2 Q53TP5_HUMAN	Q8r492 rattus norvegicus
15	2141	55.2	761	2 Q53TP5_HUMAN	Q8r492 rattus norvegicus
16	2138	55.1	761	1 SEPR_HUMAN	P12884 homo sapiens
17	2136	55.1	761	1 DPP10_XENIA	P97321 mus musculus
18	2103	54.2	755	2 Q91651_XENIA	Q91651 xenopus laevis
19	1966	50.7	860	2 Q4RQ14_TETING	Q4rcj4 tetrodon natus
20	1819	46.9	704	2 Q4S309_TETING	Q4rcj4 tetrodon natus
21	1216	31.4	796	1 DPP10_HUMAN	Q8n088 homo sapiens
22	1207	31.1	796	1 DPP10_RAT	Q6f629 rattus norvegicus
23	1192	30.8	796	2 Q6INB7_XENIA	Q6f629 rattus norvegicus
24	1182	30.5	797	1 DPP10_MOUSE	Q6nxk7 mus musculus
25	1117	28.8	865	1 DPP6_HUMAN	P42558 homo sapiens
26	1111	28.7	731	2 Q7PSF9_ANOCA	Q4rcj4 anophelles gambiae
27	1109	28.6	803	1 DPP6_PANTHERA	Q5650 pan troglodytes
28	1106	28.5	803	2 Q9QV78_IMURU	Q9QV78 rattus sp.
29	1104	28.5	859	1 DPP6_RAT	P46101 rattus norvegicus
30	1101	28.4	745	2 Q9VMN2_DRONE	P42659 drosophila melanogaster
31	1094	28.2	863	1 DPP6_BOVINE	P42659 bos taurus

ALIGNMENTS

Result 1	DPP4_HUMAN	STANDARD	PRY:	766 AA.
ID	P2487;			
AC				
DT	01-AUG-1992 (Rel. 23, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	13-SEP-2005 (Rel. 48, Last annotation update)			
DE	Dipeptidyl peptidase 4 (EC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP4)			
DE	IV (T-cell activation antigen CR26) (TPI0) (Adenosine deaminase complexing protein 2) (ADABP) [contains: Dipeptidyl peptidase 4 membrane form (Dipeptidyl peptidase IV membrane form); Dipeptidyl peptidase 4 soluble form (Dipeptidyl peptidase IV soluble form)]			
DE	Homo sapiens (Human)			
OS	Batrachotaenia; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catahrini; Hominidae; Homo.			
OC				
OC				
OC				
NCBI_TaxID	9606;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RC	TISSUE-liver;			
RC	MEDLINE=9232551; PubMed=1252704; DOI=10.1016/0167-4781(92)90036-Y;			
RA	Miwaumi Y., Hayashi Y., Arakawa F., Ichihara Y.; "Molecular cloning and sequence analysis of human dipeptidyl peptidase IV, a serine protease on the cell surface.", Biochim. Biophys. Acta 1131:333-336 (1992).			
RN	[2]			
RP	NUCLEOTIDE SEQUENCE.			
RC	TISSUE=Placenta;			
RC	MEDLINE=95012454; PubMed=7927537;			
RA	Abbott C.A., Baker E., Sutherland G.R., McCaughey G.W.; "Genomic organization, exact localization, and tissue expression of the human CD26 (dipeptidyl peptidase IV) gene.", J. Immunogenetics 21(3):313-318 (1994).			
RT	RT			
RL	Immunogenetics 40:331-338 (1994).			
RN	[3]			
RP	NUCLEOTIDE SEQUENCE			
RC	TISSUE=Peripheral blood;			
RC	MEDLINE=9232476; PubMed=13252530;			
RA	Tanaka T., Camerini D., Seed B., Torimoto Y., Dang N.H., Kameoka J., Dahlberg H.N., Schlossman S.F., Morimoto C.; "Cloning and functional expression of the T cell activation antigen CD26.", J. Immunol. 149:481-486 (1992).			
RN	[4]			
RP	ERATUM.			
RX	MEDLINE=93171637; PubMed=8094732;			
RA	Tanaka T.; Dahlberg H.N., Schlossman S.F., Morimoto C.; "Cloning and functional expression of the T cell activation antigen CD26.", J. Immunol. 150:2090-2090 (1993).			
RN	[5]			
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].			
RC	TISSUE=Prostate, and Uterus;			
RX	MEDLINE=22388257; PubMed=1247932; DOI=10.1073/pnas.242603899;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Klauser R.D., Colling F.S., Wagner L., Shenmen C.M., Schuler G.D.,			

- RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L., Scheetz T.E., RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E., RA Brownstein M.J., Soares M.B., Bonaldo M.P., Casavant T.L., Prange C., Carinucci P., Mullany S.J., Toshiyuki S., Loquellano P.N., Peters G.J., Abramson R.D., Mullany S.J., RA Boak S.A., McBwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.W., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., RA Villalobos D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., FAhey J., Heitton B., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmitz J., Myers R.M., RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E., Scherich A., Schein J.E., Jones S.J.M., Marra M.A.; "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences"; Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002). [6]
- RN NUCLEOTIDE SEQUENCE OF 1-551.
- RP TISSUE=Colon; MEDLINE=2165847; PubMed=1347043;
- RC TISSUE=Colon; MEDLINE=2165847; PubMed=1347043;
- RA Darmoul D., Lacasa M., Baricault L., Marguet D., Sapin C., Trotot P., Barbat A.; "Dipeptidyl peptidase IV (CD 26) gene expression in enterocyte-like colon cancer cell lines HT-29 and Caco-2. Cloning of the complete human coding sequence and changes of dipeptidyl peptidase IV mRNA levels during cell differentiation.", J. Biol. Chem. 267:4824-4833(1992). [7]
- RN NUCLEOTIDE SEQUENCE OF 545-766.
- RC TISSUE=Colon; MEDLINE=91024044; PubMed=1977364;
- RA Darmoul D., Lacasa M., Chauvet I., Swallow D., Trugnan G.; "Isolation of a cDNA probe for the human intestinal dipeptidylpeptidase IV and assignment of the gene locus DPP4 to chromosome 2.", Ann. Hum. Genet. 54:191-197(1990). [8]
- RN NUCLEOTIDE SEQUENCE OF 1-31.
- RP TISSUE=Kidney; MEDLINE=16067599; PubMed=7487939;
- RC TISSUE=Kidney; MEDLINE=16067599; PubMed=7487939;
- RA Boehm S.K., Gurn J.R.Jr., Erickson R.H., Hicks J.W., Kim Y.S.; RA Marouf S.; "Human dipeptidyl peptidase IV gene promoter: tissue-specific regulation from a TATA-less GC-rich sequence characteristic of a housekeeping gene promoter"; Biochem. J. 311:835-843(1995). [9]
- RN PROTEIN SEQUENCE OF 1-22, AND TISSUE SPECIFICITY.
- RP TISSUE=Kidney; PubMed=1677636;
- RC TISSUE=Kidney; MEDLINE=1677636; PubMed=1677636;
- RA Gorvel J.P., Ferrero A., Chambraud L., Rigal A., Bonicel J., Marouf S.; "Expression of sucrase-isomaltase and dipeptidylpeptidase IV in human small intestine and colon"; Gastroenterology 101:618-625(1991). [10]
- RN PARTIAL PROTEIN SEQUENCE.
- RP TISSUE=Kidney; MEDLINE=93210468; PubMed=8036237; DOI=10.1084/jem.177.4.1135;
- RC TISSUE=Kidney; MEDLINE=93210468; PubMed=8036237; DOI=10.1084/jem.177.4.1135;
- RA Morrison M.E., Vijayasaradhi S., Engelstein D., Albino A.P., Houghton A.N.; "A marker for neoplastic progression of human melanocytes is a cell surface ectopeptidase"; J. Exp. Med. 177:1135-1143(1993). [11]
- RN X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 38-766 IN COMPLEX WITH INHIBITOR, AND HOMODIMERIZATION.
- RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 38-766 IN COMPLEX WITH INHIBITOR, AND HOMODIMERIZATION.
- RA Houghton A.N.; "The structure and function of human dipeptidyl peptidase IV, possessing a unique eight-bladed beta-propeller fold."; RA Richards S., D'Arcy A., Mac Sweeney A., Pierau S., Gardiner R., Dale G.E.; "High-resolution structure of human apo dipeptidyl peptidase IV/CD26 and its complex with 1-((2-(5'-odopyridin-2-yl)amino)-ethyl)aminocetyl acetyl-2-cyano-(S)-pyrrolidine"; Acta Crystallogr. D 59:1206-1212(2003). [12]
- RA X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 35-771, HOMODIMERIZATION, AND N-GLYCOSYLATION SITES. PubMed=1264248; DOI=10.1016/S0006-291X(03)00258-4;
- RC X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 35-771, HOMODIMERIZATION, AND N-GLYCOSYLATION SITES. PubMed=1264248; DOI=10.1016/S0006-291X(03)00258-4;
- RA Hiramatsu H., Kyono K., Higashiyama Y., Fukushima C., Shima H., Sugiyama S., Inaka K., Yamamoto A.; "Shimizu R.; "The structure and function of human dipeptidyl peptidase IV, possessing a unique eight-bladed beta-propeller fold."; RA Richards S., D'Arcy A., Mac Sweeney A., Pierau S., Gardiner R., Dale G.E.; "High-resolution structure of human apo dipeptidyl peptidase IV/CD26 and its complex with 1-((2-(5'-odopyridin-2-yl)amino)-ethyl)aminocetyl acetyl-2-cyano-(S)-pyrrolidine"; Acta Crystallogr. D 59:1206-1212(2003). [13]
- RA X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 39-766, HOMODIMERIZATION, AND N-GLYCOSYLATION SITES. PubMed=12483204; DOI=10.1016/nab882;
- RC X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 39-766, HOMODIMERIZATION, AND N-GLYCOSYLATION SITES. PubMed=12483204; DOI=10.1016/nab882;
- RA Pasmusen H.B., Branner S., Viberg F.C., Wagmann N.; "Crystal structure of human dipeptidyl peptidase IV/CD26 in complex with a substrate analog"; Nat. Struct. Biol. 10:19-25(2003). [14]
- RA X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 39-766, HOMODIMERIZATION, AND N-GLYCOSYLATION SITES. PubMed=12206826; DOI=10.1016/S0969-2126(03)00160-6;
- RC X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 39-766, HOMODIMERIZATION, AND N-GLYCOSYLATION SITES. PubMed=12206826; DOI=10.1016/S0969-2126(03)00160-6;
- RA Thoma R., Loeffler B., Stihle M., Huber W., Ruf A., Henning M.; "Structural basis of proline-specific exopeptidase activity as observed in human dipeptidyl peptidase-IV"; Structure 11:947-959(2003). [15]
- RA -1- FUNCTION: Removes N-terminal dipeptides sequentially from polypeptides having unsubstituted N-terminal provided that the penultimate residue is proline. Plays a role in T cell activation. CC -1- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-[Zaa], from a polypeptide, preferentially when Yaa is Pro, provided CC -1- SUBUNIT: Homodimer or heterodimer with Seprase (FAP).
- RA -1- SUBCELLULAR LOCATION: Type II membrane protein. Also exists in a soluble form. CC -1- TISSUE SPECIFICITY: Expressed in the poorly differentiated crypt cells of the small intestine as well as in the mature villous cells. Expressed at very low levels in the colon. CC -1- PTM: The soluble form (SDPP) derives from the membrane form (MDPP) by proteolytic processing. CC -1- SIMILARITY: Belongs to the peptidase S9B family. DPP4 subfamily. CC -1- This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed. CC -1- DR EMBL; U13735; AAB60646; 1 ; JOINED; Genomic_DNA. DR EMBL; U13710; AAB60646; 1 ; JOINED; Genomic_DNA. DR EMBL; U13711; AAB60646; 1 ; JOINED; Genomic_DNA. DR EMBL; U13712; AAB60646; 1 ; JOINED; Genomic_DNA. DR EMBL; U13713; AAB60646; 1 ; JOINED; Genomic_DNA. DR EMBL; U13714; AAB60646; 1 ; JOINED; Genomic_DNA. DR EMBL; U13715; AAB60646; 1 ; JOINED; Genomic_DNA. DR EMBL; U13716; AAB60646; 1 ; JOINED; Genomic_DNA. DR EMBL; U13717; AAB60646; 1 ; JOINED; Genomic_DNA. DR EMBL; U13718; AAB60646; 1 ; JOINED; Genomic_DNA. DR EMBL; U13719; AAB60646; 1 ; JOINED; Genomic_DNA. DR EMBL; U13720; AAB60646; 1 ; JOINED; Genomic_DNA. DR EMBL; U13721; AAB60646; 1 ; JOINED; Genomic_DNA. DR EMBL; U13722; AAB60646; 1 ; JOINED; Genomic_DNA. DR EMBL; U13723; AAB60646; 1 ; JOINED; Genomic_DNA. DR EMBL; U13724; AAB60646; 1 ; JOINED; Genomic_DNA. DR EMBL; U13725; AAB60646; 1 ; JOINED; Genomic_DNA. DR EMBL; U13726; AAB60646; 1 ; JOINED; Genomic_DNA. DR EMBL; U13727; AAB60646; 1 ; JOINED; Genomic_DNA. DR EMBL; U13728; AAB60646; 1 ; JOINED; Genomic_DNA. DR EMBL; U13729; AAB60646; 1 ; JOINED; Genomic_DNA. DR EMBL; U13730; AAB60646; 1 ; JOINED; Genomic_DNA. DR EMBL; U13731; AAB60646; 1 ; JOINED; Genomic_DNA. DR EMBL; U13732; AAB60646; 1 ; JOINED; Genomic_DNA. DR EMBL; U13733; AAB60646; 1 ; JOINED; Genomic_DNA. DR EMBL; U13734; AAB60646; 1 ; JOINED; Genomic_DNA.

DR	EMBL; M74777; AAAS1943.1; - ; mRNA.		
DR	EMBL; BC013329; AAH13329.2; - ; mRNA.		
Query Match	Score 3877; DB 1; Length 766;		
Best Local Similarity	100.0%; Pred. No. 3..6e-256;		
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1 NTYRLKLYSTRLWISDHEYLKOENNVLVFNAYEGNNSVLENSTDEFGHSINDYSISPD 60		
Db	51 NTYRLKLYSTRLWISDHEYLQKENNLVFNAYEGNNSVLENSTDEFGHSINDYSISPD 110		
Qy	61 QFILLENTYKQWRHISYTASYDIVLNKLQLITEERIPNTQWTVSPYGHKLAYWN 120		
Db	111 QFILLENTYKQWRHISYTASYDIVLNKLQLITEERIPNTQWTVSPYGHKLAYWN 170		
Qy	121 DIFYKIEPNLPSYRITWTGKEDITINGITDMVYEEFSYASALWSPNCTFLAYAQFND 180		
Db	171 DIFYKIEPNLPSYRITWTGKEDITINGITDMVYEEFSYASALWSPNCTFLAYAQFND 230		
Qy	181 TEVPLIEYSFSDESQYKPTKTRVPYPKAGAVNPYTFKFVYNTDSLSVTNATSQTATP 240		
Db	231 TEVPLIEYSFSDESQYKPTKTRVPYPKAGAVNPYTFKFVYNTDSLSVTNATSQTATP 290		
Qy	241 ASMLIGDHYLCDVTWATQERISLQWLRIONYQSYNDICDYESESRWNLCLVAROHIESTM 300		
Db	291 ASMLIGDHYLCDVTWATQERISLQWLRIONYQSYNDICDYESESRWNLCLVAROHIESTM 350		
Qy	301 TGWVGRFRPSRSPHETLDGSNSYKTSISNEEGYRHYCIFQIDKKDTFTTGTWEVIGEAL 360		
Db	351 TGWVGRFRPSRSPHETLDGSNSYKTSISNEEGYRHYCIFQIDKKDTFTTGTWEVIGEAL 410		
Qy	361 TSDYLIYISNEYKGMPGRNLKYKQSLDVTKTCSCELNPEROCCYSSFSKEAKYIQL 420		
Db	411 TSDYLIYISNEYKGMPGRNLKYKQSLDVTKTCSCELNPEROCCYSSFSKEAKYIQL 470		
Qy	421 RCGSGPGLPLYLHSSYNDKGRLVLDNSALDKMLQNQVMSKQLDFIILNETKFWYQML 480		
Db	471 RCGSGPGLPLYLHSSYNDKGRLVLDNSALDKMLQNQVMSKQLDFIILNETKFWYQML 530		
Qy	481 PPHEFDKSKYKTPLLDQYAGPCSQKADTVFLNWLQYAGPCSQKADTVFLNWLQYQGKCI 540		
Db	531 PPHEFDKSKYKTPLLDQYAGPCSQKADTVFLNWLQYAGPCSQKADTVFLNWLQYQGKCI 590		
Qy	541 MHAINRRLGTFEVEDQIYEARQFSMGMFDNKRAIWGNSYGGUTSMVLGSGGVFKCG 600		
Db	591 MHAINRRLGTFEVEDQIYEARQFSMGMFDNKRAIWGNSYGGUTSMVLGSGGVFKCG 650		
Qy	601 IAVAPVSREWWYDSYTERTMGLPPTPEDNDHYRNSTNSRAENFKQVEYLILLHTADDN 660		
Db	651 IAVAPVSREWWYDSYTERTMGLPPTPEDNDHYRNSTNSRAENFKQVEYLILLHTADDN 710		
Qy	661 VHFQOSAQTSKALVDVGVDFOAMMYTDEDHGIAASSTAHOHYTMSHFKQCFSLP 716		
Db	711 VHFQOSAQTSKALVDVGVDFOAMMYTDEDHGIAASSTAHOHYTMSHFKQCFSLP 766		
Qy	766 AA.		
RESULT 2			
ID	Q53TN1_HUMAN PRELIMINARY;		
AC	Q53TN1;		
DT	13-SEP-2005 (Tremblrel. 31, Created)		
DT	13-SEP-2005 (Tremblrel. 31, Last sequence update)		
DT	13-SEP-2005 (Tremblrel. 31, Last annotation update)		
DB	Hypothetical protein DPP4.		
GN	Name=DPP4;		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Catarrhini; Hominidae;		
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Sapetti L.; Homo.		
OX	[1]		
RN	NCBI_TaxID=9606;		
RP	NUCLEOTIDE SEQUENCE.		
RA	Ozanich A., Stonerking T., Hawkins M., Sapetti L.;		
Qy	541 MHAINRRLGTFEVDQIYEARQFSMGMFDNKRAIWGNSYGGUTSMVLGSGGVFKCG 600		

Db	591	MHAINRRLGTPFEVEDQIEARQFSKMGFVDNKR1AIWGMSYGGVTSMLGSSGVFFKG	650	Qy	301	TGMYGRFRPSEPHETLDGNSFYK1ISNEEGYRHICYFOIDKKDCTPITKGTMEVIGTEAL	360
Qy	601	IAVAPVSRNEYDSVYTERMGLPTPEVDLHYRNSTMSRAENFKOQEYLILHGTA	660	Db	350	TGMYGRPSEPHETLDGNSFYK1ISNEEGYRHICYFOIDKKDCTPITKGTMEVIGTEAL	409
Db	651	IAVAPVSRNEYDSVYTERMGLPTPEVDLHYRNSTMSRAENFKOQEYLILHGTA	710	Qy	361	TSDLYLTTISNEYKMPGCRNLXK1QLSDYTKTCSCLPNPERCQYVSFSKEAKYQOL	420
Qy	661	VHQOSAQ1SKALVQDVGDFQAMWYTFDEHDGIASTAHQ1YTHMSHFKQCFLP	716	Db	410	TSDLYLTTISNEYKMPGCRNLXK1QLSDYTKTCSCLPNPERCQYVSFSKEAKYQOL	469
Db	711	VHQOSAQ1SKALVQDVGDFQAMWYTFDEHDGIASTAHQ1YTHMSHFKQCFLP	766	Qy	421	RCSGPGLPLYTHS SYNDKGLEVLNEDSALDKMLQNYQMPSKCLDFILNETKFQYQML	480
RESULT 3							
Q5RTG7	PONPY	PONPY PRELIMINARY;	PRT;	765	AA.	Qy	481
ID	Q5RTG7-					PPHFDSKSKYKPLDLYAGPCSKADTYVFLRNLWATYLASTENIIVASPDGRGSGYQGDKI	540
AC	Q5RTG7-					PPHFDSKSKYKPLDLYAGPCSKADTYVFLRNLWATYLASTENIIVASPDGRGSGYQGDKI	540
DT	01-FEB-2005	(TREMBIrel. 29, Created)				PPHFDSKSKYKPLDLYAGPCSKADTYVFLRNLWATYLASTENIIVASPDGRGSGYQGDKI	589
DT	01-FEB-2005	(TREMBIrel. 29, Last sequence update)				PPHFDSKSKYKPLDLYAGPCSKADTYVFLRNLWATYLASTENIIVASPDGRGSGYQGDKI	600
DE	Hypothetical protein DKFPz469P1419.					PPHFDSKSKYKPLDLYAGPCSKADTYVFLRNLWATYLASTENIIVASPDGRGSGYQGDKI	649
GN	Name=DKFPz469P1419;					PPHFDSKSKYKPLDLYAGPCSKADTYVFLRNLWATYLASTENIIVASPDGRGSGYQGDKI	649
OS	Pongo pygmaeus (Orangutan)					PPHFDSKSKYKPLDLYAGPCSKADTYVFLRNLWATYLASTENIIVASPDGRGSGYQGDKI	660
OC	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					PPHFDSKSKYKPLDLYAGPCSKADTYVFLRNLWATYLASTENIIVASPDGRGSGYQGDKI	660
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;					PPHFDSKSKYKPLDLYAGPCSKADTYVFLRNLWATYLASTENIIVASPDGRGSGYQGDKI	660
OC	Pongo.					PPHFDSKSKYKPLDLYAGPCSKADTYVFLRNLWATYLASTENIIVASPDGRGSGYQGDKI	660
NCBI_TaxID=9600;						PPHFDSKSKYKPLDLYAGPCSKADTYVFLRNLWATYLASTENIIVASPDGRGSGYQGDKI	660
RN	[1]					PPHFDSKSKYKPLDLYAGPCSKADTYVFLRNLWATYLASTENIIVASPDGRGSGYQGDKI	660
RP	NUCLEOTIDE SEQUENCE.					PPHFDSKSKYKPLDLYAGPCSKADTYVFLRNLWATYLASTENIIVASPDGRGSGYQGDKI	660
RC	TISSUE=Kidney;					PPHFDSKSKYKPLDLYAGPCSKADTYVFLRNLWATYLASTENIIVASPDGRGSGYQGDKI	660
RG	The German cDNA Consortium;					PPHFDSKSKYKPLDLYAGPCSKADTYVFLRNLWATYLASTENIIVASPDGRGSGYQGDKI	660
RA	Poustka A., Albert R., Moosmayer P., Schupp I., Wellenreuther R.,					PPHFDSKSKYKPLDLYAGPCSKADTYVFLRNLWATYLASTENIIVASPDGRGSGYQGDKI	660
RA	Meves H.W., Weil B., Amid C., Obanger A., Fobo G., Han M., Wiemann S.;					PPHFDSKSKYKPLDLYAGPCSKADTYVFLRNLWATYLASTENIIVASPDGRGSGYQGDKI	660
RL	Submitted (Nov-2004) to the EMBL/GenBank/DBJ databases.					PPHFDSKSKYKPLDLYAGPCSKADTYVFLRNLWATYLASTENIIVASPDGRGSGYQGDKI	660
DR	EMBL: CB861050; CATH92293.1; -; mRNA.					PPHFDSKSKYKPLDLYAGPCSKADTYVFLRNLWATYLASTENIIVASPDGRGSGYQGDKI	660
DR	SRM, Q5RTG7, 38-765.					PPHFDSKSKYKPLDLYAGPCSKADTYVFLRNLWATYLASTENIIVASPDGRGSGYQGDKI	660
DR	GO: GO:0016040; C membrane; IEA.					PPHFDSKSKYKPLDLYAGPCSKADTYVFLRNLWATYLASTENIIVASPDGRGSGYQGDKI	660
DR	GO: GO:0004177; F amineopeptidase activity; IEA.					PPHFDSKSKYKPLDLYAGPCSKADTYVFLRNLWATYLASTENIIVASPDGRGSGYQGDKI	660
DR	GO: GO:0004274; F dipeptidyl-peptidase IV activity; IEA.					PPHFDSKSKYKPLDLYAGPCSKADTYVFLRNLWATYLASTENIIVASPDGRGSGYQGDKI	660
DR	GO: GO:0016787; F hydrolase activity; IEA.					PPHFDSKSKYKPLDLYAGPCSKADTYVFLRNLWATYLASTENIIVASPDGRGSGYQGDKI	660
DR	GO: GO:0004287; F prolyl oligopeptidase activity; IEA.					PPHFDSKSKYKPLDLYAGPCSKADTYVFLRNLWATYLASTENIIVASPDGRGSGYQGDKI	660
DR	GO: GO:0006508; P proteolysis and peptidolysis; IEA.					PPHFDSKSKYKPLDLYAGPCSKADTYVFLRNLWATYLASTENIIVASPDGRGSGYQGDKI	660
DR	InterPro: IPR002471; Pept_S9_AS.					PPHFDSKSKYKPLDLYAGPCSKADTYVFLRNLWATYLASTENIIVASPDGRGSGYQGDKI	660
DR	InterPro: IPR001375; Pept_IDAE_S9.					PPHFDSKSKYKPLDLYAGPCSKADTYVFLRNLWATYLASTENIIVASPDGRGSGYQGDKI	660
DR	InterPro: IPR002469; Peptidase_S9B.					PPHFDSKSKYKPLDLYAGPCSKADTYVFLRNLWATYLASTENIIVASPDGRGSGYQGDKI	660
DR	InterPro: IPR000319; Ser_enzrs.					PPHFDSKSKYKPLDLYAGPCSKADTYVFLRNLWATYLASTENIIVASPDGRGSGYQGDKI	660
DR	Pfam: PF00930; DPPY_N; 1.					PPHFDSKSKYKPLDLYAGPCSKADTYVFLRNLWATYLASTENIIVASPDGRGSGYQGDKI	660
DR	PFEM; PF00326; Peptidase_S9; 1.					PPHFDSKSKYKPLDLYAGPCSKADTYVFLRNLWATYLASTENIIVASPDGRGSGYQGDKI	660
KW	Hydrolease; Hypothetical protein; Transmembrane.					PPHFDSKSKYKPLDLYAGPCSKADTYVFLRNLWATYLASTENIIVASPDGRGSGYQGDKI	660
SEQ	SEQUENCE 765 AA; 88038 MW;					PPHFDSKSKYKPLDLYAGPCSKADTYVFLRNLWATYLASTENIIVASPDGRGSGYQGDKI	660
Query Match							
Best Local Similarity	99.3%	Score 3848;	DB:	2	Length 765;	RP	NUCLEOTIDE SEQUENCE.
Matches	710;	Conservative	2;	Mismatches	4;	RC	TISSUE=Lymphocyte;
						RX	PubMed=22067734; PubMed=12073152; DOI=10.1007/s00251-002-0456-6;
Qy	1	NTYRLKLISLRWISDHKEYQKENNLVNAEYGNSSPLENSTEDFGHSINDYSSPD	60	RA	Lee S.-U., Park Y.-H., Davis W.C., Hamilton M.J., Naessens J., Bohach G.A.;		
Db	50	NTYRLKLISLRWISDHKEYQKENNLVNAEYGNSSPLENSTEDFGHSINDYSSPD	109	RA	"Molecular characterization of bovine CD26 upregulated by a staphylococcal superantigen."		
Qy	61	GQFILEENYVKWKRHSSTASYDIYLNKRQLITEERLPNNNTOWTWSPVGHKLAYWN	120	RT	Eur. J. Immunol. 34:216-220(2002).		
Db	110	GQFILEENYVKWKRHSSTASYDIYLNKRQLITEDRPNNTOWTWSPVGHKLAYWN	169	RT	PROTEIN SEQUENCE OF 1-24.		
Qy	121	DIYKIEPNPLPSRITWTCKEDITYGNTDWYEEVEFSAYSALWWSPNGTFLAYAOFND	180	RC	TISSUE=Thymus;		
Db	170	DIVKIEPALPSRITWTCKEDITYGNTDWYEEVEFSAYSALWWSPNGTFLAYAOFND	229	RA	Medline=22021197; PubMed=11981836;		
Qy	181	TEVPLIEFSYFSDESQPKTRVPRVPKAGAVNPTRKEFVNNTDSLSSVTNATSIQTAP	240	RA	Gliddon D.R., Howard C.J., Lee S.-U., Park Y.-H., Davis W.C., Hamilton M.J., Naessens J.,		
Db	230	TEVPLIEFSYFSDESQPKTRVPRVPKAGAVNPTRKEFVNNTDSLSSVTNATSIQTAP	289	RT	"CD26 is expressed on a restricted subpopulation of dendritic cells in		
Qy	241	ASMLIGDHYLCDVTVATWERTISLOWLRQINTSYMDICDYDESSGRNCLVARQHTEMST	300	RA	Lee S.-U., Perens W., Davis W.C., Hamilton M.J., Naessens J.,		
Db	290	ASMLIGDHYLCDVTVATWERTISLOWLRQINTSYMDICDYDESSGRNCLVARQHTEMST	349	RA	Fox L.K.,		

RA Naessens J., Bohach G.A.; "Identity of activation molecule 3 on superantigen-stimulated bovine RT. cells is CD26." *Infect. Immun.* 69:7190-7193 (2001).
 RT. [4]
 RN PROTEIN SEQUENCE OF 537-546.
 RP TISSUE=Kidney;
 RC MEDLINE=98293306; PubMed=9629661; DOI=10.1016/S0305-0491(97)00327-1;
 RA Ben-Shooshan I.; Parola A.H.;
 RT "The CP-I subunit of adenosine deaminase complexing protein from calf IV."
 RT kidney is identical to human, mouse, and rat dipeptidyl peptidase
 RL Comp. Biochem. Physiol. 119B:289-292 (1998).
 CC -I- FUNCTION: Removes N-terminal dipeptides sequentially from
 CC polypeptides having unsubstituted N-termini. Provided that the
 CC penultimate residue is proline. Binds and regulates the activity
 CC of ADA.
 CC -I- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-[
 CC zaa], from a polypeptide, preferentially when Yaa is Pro, provided
 CC that Zaa is neither Pro nor hydroxyproline.
 CC -I- SUBUNIT: Homodimer.
 CC -I- SUBCELLULAR LOCATION: Type II membrane protein. Also exists in a
 CC soluble form (By similarity).
 CC -I- TISSUE SPECIFICITY: Intestinal epithelium, dendritic cells and
 CC several immune system tissues.
 CC -I- PM: The soluble form (SDPP) derives from the membrane form (MDPP)
 CC by proteolytic processing (By similarity).
 CC -I- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 DR EMBL; AF0161805; AAU67836.1; mRNA.
 DR HSSP; AV056834; AAU23628.1; mRNA.
 DR P27487; 1PFO.
 DR SMR; PB1425; 38-764.
 DR MEROPS; S09.003/-.
 DR InterPro; IPR00471; Pept_S9_AS.
 DR InterPro; IPR001375; Peptidase_S9.
 DR InterPro; IPR002469; Peptidase_S9B.
 DR InterPro; IPR000379; Ser_ester.
 DR Pfam; PF00930; DPPIV_N_1.
 DR PROSITE; PS00326; Peptidase_S9_1.
 KW Aminopeptidase; Direct protein sequencing; Glycoprotein; Hydrolase;
 KW Protease; Serine protease; Signal-anchor; Transmembrane.
 FT CHAIN 1 765 Dipeptidyl peptidase 4 membrane form.
 FT CHAIN 38 765 Dipeptidyl peptidase 4 soluble form (By similarity).
 FT TOPO_DOM 1 6 Cytoplasmic (Potential).
 FT TRANSMEM 7 29 Signal-anchor for type II membrane protein (Potential).
 FT ACT_SITE 30 765 Extracellular (Potential).
 FT ACT_SITE 629 629 Charge relay system (By similarity).
 FT ACT_SITE 707 707 Charge relay system (By similarity).
 FT ACT_SITE 739 739 N-linked (GlcNAc. . .) (By similarity).
 FT CARBOHYD 84 84 N-linked (GlcNAc. . .) (By similarity).
 FT CARBOHYD 91 91 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 149 149 N-linked (GlcNAc. . .) (By similarity).
 FT CARBOHYD 218 218 N-linked (GlcNAc. . .) (By similarity).
 FT CARBOHYD 228 228 N-linked (GlcNAc. . .) (By similarity).
 FT CARBOHYD 271 271 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 280 280 N-linked (GlcNAc. . .) (By similarity).
 FT CARBOHYD 320 320 N-linked (GlcNAc. . .) (By similarity).
 FT CARBOHYD 392 392 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 495 495 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 684 684 N-linked (GlcNAc. . .) (By similarity).
 FT DISULFID 384 393 By similarity.
 FT DISULFID 443 446 By similarity.
 FT DISULFID 453 471 By similarity.

	PT SQ	DISULFID	648	761 AA;	765 AA;	88369 MW;	E32165421F43E116 CRC64;
Query Match	91.1%	Score 3531;	DB 1;	Length 765;			
Best local Similarity	89.4%		Pred. No. 1.	1.6e-232	Gaps	0;	
Matches	639;	Conservative	40;	Mismatches	36;	Indels	0;
1	NTYFLKLYSLRWSDEHELYKQENNLVYNAYGNSSVYLENSTDFEFSHSINDYSISPD 60						
50	NTFFMKFTNLRWVSDHELYKQENNLVYNAYGNSSVYLENSTDFEFSHSINDYSISPD 109						
61	GQFILLETNVKQRHHTASYDIDLNKQLITEERIPNNNTOWTVTWSVGHKLAYWN 120						
110	RQYLFETNVKQRHHTASYDIDLNKQLITEERIPNNNTOWTVTWSVGHKLAYWN 169						
121	DIYKTIKEPLNPLSPTRITWGKDIDTYINGTDWYEEVEVSAYSAWWSPNTLFLAYQFND 180						
170	DIYKNEPNSPSPSQTWTGKDVYINGTDWYEEVEVSAYSAWWSPNTLFLAYQFND 229						
181	TEVPLIEFSYSDSLOPQTKTVPYPIKAGANPNTVKPFVNNTDSLSSVTNATSIQTAP 240						
230	TEVPLIEFSYSDSLOPQTKTVPYPIKAGANPNTVKPFVNNTDSLSSVTNATSIQTAP 289						
241	ASMLJGDHYLCDVWTATOBRISLQWLRLRQNYSVMDICDDESSGRMNLVARHIENST 300						
290	GSVLJGDHYLCDVWTWTERISLQWLRLRQNYSVMDICDYDRSTGRWISSVGRHIBIST 349						
301	TGNYGRFREPSEPHTLGDSNFSYKILISNEGYRICYFQDKDCTPFTKGTMWVIGTEAL 360						
350	TGWWGRFRREPHTSDGNSFSYKILISNEGYKHHIFDFTKRNTFTKGAWEVIGTEAL 409						
361	TSDLYYISNEYKEMPGRNLKYKQLSDYTKTCLSCSLNPERQYYSVSFSKEAKYQOL 420						
410	TSDDLYYISNEYKEMPGRNLKYKQLSDYTKTCLSCSLNPERQYYSVSFSKEAKYQOL 459						
421	RCSSPGLPDTYUHSISVNDKGLRVLEDNSALDKRMQLNDYTKTCLSCSLNPERQYYSVSFSKEAKYQOL 480						
470	RCSSPGLPDTYUHSISVNDKGLRVLEDNSALDKRMQLNDYTKTCLSCSLNPERQYYSVSFSKEAKYQOL 529						
481	PPHDKSKKYKPLIILDVYAGPCSKOKADTFLRNLWATYLASTENITIVASPDGRGSCYQGDKI 540						
530	PPHDKSKKYKPLIILDVYAGPCSKOKADTFLRNLWATYLASTENITIVASPDGRGSCYQGDKI 589						
541	MHAINRRIGTGFVBDQIAARQSSKMGTVDNKRIAIWGMYSGGVTSMLGCSGGVPGCG 600						
590	MHAINRRIGTGFVBDQIAARQSSKMGTVDNKRIAIWGMYSGGVTSMLGCSGGVPGCG 649						
601	IAVPVPSWEYDSSVYTERYMGILTPEDNLDHGTADN 660						
650	IAVPVPSWEYDSSVYTERYMGILTPEDNLDSTMSRNSTMRAENFKQVEYLHGTADN 709						
661	VHEQSAQISKALVDVGIDFQAMWYTDDEDHGIASTTAHQHITHMSHTIKOCFL 715						
710	VHEQSAQISKALVDAGYDFQSWMYTDDEDHGIASTTAHQHITHMSHTIKOCFL 764						
RESULT 5							
DPP4_PELCA	STANDARD;	PRT;	765 AA.				
ID	DPP4_PELCA						
AC	Q92II7;						
DT	28-FEB-2003	(Rel. 41, Created)					
DT	28-FEB-2003	(Rel. 41, Last sequence update)					
DT	13-SEP-2005	(Rel. 48, Last annotation update)					
DE	Dipeptidyl peptidase 4 (EC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP IV) ("r-cell activation antigen CD26)" [Contigs: Dipeptidyl peptidase 4, Membrane form (Dipeptidyl peptidase IV membrane form); Dipeptidyl peptidase 4 soluble form (Dipeptidyl peptidase IV soluble form)]. Name=DPP4; Synonyms=CD26; Cat).						
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Butelostomi; Mammalia; Eutheria; Laurasiatheria; Carnivora; Pissipedia; Felidae; Felinae; Felis; Cat.						
OC	Ncbi_TaxID=9885;						

RN	[1] NUCLEOTIDE SEQUENCE.
RC	TISSUE_Periphera_1 blood;
RX	MEDLINE=20094000; PubMed=10630304; DOI=10.1007/s002510050616;
RA	Nishimura Y., Miyazawa T., Ikeda Y., Izumiya Y., Nakamura K., Sato E., Takahashi E.; "Molecular cloning and sequencing of a cDNA encoding the feline T-cell activation antigen CD26 homologue.";
RU	Immunogenetics 50:366-368 (1999).
CC	-1 - FUNCTION: Removes N-terminal dipeptides sequentially from polypeptides having unsubstituted N-terminal provided that the penultimate residue is proline (By similarity).
CC	-1 - CARAUTRIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-[-zaa-], from a polypeptide, preferentially when Yaa is Pro, provided Zaa is neither Pro nor hydroxyproline.
CC	-1 - SUBUNIT: Homodimer (By similarity).
CC	-1 - SUBCELLULAR LOCATION: Type II membrane protein. Also exists in a soluble form (By similarity).
CC	-1 - PTM: The soluble form (SDPP) derives from the membrane form (MDPP) by proteolytic processing (By similarity).
CC	-1 - SIMILARITY: Belongs to the Peptidase S9B family. DPPIV subfamily.
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.
CC	-----
DR	AB023952; BA92344.1; - ; mRNA.
DR	HSSP; P27487; 1PFQ.
DR	SMK; Q9N217; 38-765.
DR	MEIOPS; S09_003; InterPro; IPR00471; Pept_S9_AS.
DR	InterPro; IPR01315; Peptidae_S9.
DR	InterPro; IPR002459; Peptidae_S9B.
DR	Pfam; PF00930; DPP4_N_1.
DR	PROSITE; PS00708; PRO_ENDOPP_1.
KW	Aminopeptidase; Glycoprotein; Hydrolase; Protease; Serine protease; Signal-anchor; Transmembrane.
PT	CHAIN 1 765 Peptidase 4 membrane form.
PT	CHAIN 38 765 Peptidase 4 soluble form (By similarity).
PT	TOPO_DOM 1 6 Cytoplasmic (Potential).
PT	TRANSMEM 7 29 Signal-anchor for type II membrane protein (Potential).
PT	TOPO_DOM 30 765 Extracellular (Potential).
PT	ACT_SITE 629 629 Charge relay system (By similarity).
PT	ACT_SITE 707 707 Charge relay system (By similarity).
PT	ACT_SITE 739 739 Charge relay system (By similarity).
PT	CARBHYD 84 84 N-linked (GlcNAc. . .) (By similarity).
PT	CARBHYD 91 91 N-linked (GlcNAc. . .) (By similarity).
PT	CARBHYD 149 149 N-linked (GlcNAc. . .) (By similarity).
PT	CARBHYD 178 178 N-linked (GlcNAc. . .) (Potential).
PT	CARBHYD 228 228 N-linked (GlcNAc. . .) (By similarity).
PT	CARBHYD 280 280 N-linked (GlcNAc. . .) (By similarity).
PT	CARBHYD 320 320 N-linked (GlcNAc. . .) (By similarity).
PT	CARBHYD 330 330 N-linked (GlcNAc. . .) (Potential).
PT	CARBHYD 331 331 N-linked (GlcNAc. . .) (Potential).
PT	CARBHYD 519 519 N-linked (GlcNAc. . .) (By similarity).
PT	CARBHYD 684 684 N-linked (GlcNAc. . .) (By similarity).
PT	DISULFID 384 393 By similarity.
PT	DISULFID 443 446 By similarity.
PT	DISULFID 453 471 By similarity.
PT	DISULFID 648 761 By similarity.
SEQUENCE	765 AA; 88213 MW; 3EPCB98A22B175D9 CRC64;
SQ	Query Match 90.3%; Score 3501; DB 1; Length 765;
	Best Local Similarity 88.0%; Pred. No. 1.86-230;
	Matches 630; Conservative 46; Mismatches 40; Indels 0; Gaps 0;

1 NYTFLKLYSLRWISDHEVLYKQENNLVFNAEYGNSSUPLNSTFDEFGHSINDYSISPD 60

QY

Db	50 NTFRVKCYSLRWISDHEVLYKQDNNTLFNAEYGNSSUPLNSTFDEFGHSINDYSISPD 109
Qy	61 GQFILLEYNYKQRWRSYTAVIDYDINKQLITEERIPANTQWVTPVGHLAYWN 120
Db	110 GQFILLEYNYKQRWRSYTAVIDYDINKQLITEERIPANTQWVTPVGHLAYWN 169
Qy	121 DYYVKIEPNLPSYRWTGKEDLYNGITDANVYEEVFSAYSALWNSPNGTFLAYAOFND 180
Db	170 DYVKNBENSSSHRWTGGENAYINGADWYEETPSAYSALWNSPNGTFLAYAOFND 229
Qy	181 TEVPLIEFSYPSDESLOYPKTVRYPKAGAVNPYKFEVNTIDSSLSSVTNTSQTAP 240
Db	230 TQVPLIEFSYPSDESLOYPKTMKRPYPKAGAANPVTFLVKTDLNPNTNTSVEITPP 289
Qy	241 ASML1GPHYLCDVWATQERTISLQWLRLRIQYNSVMDICDYDESSGRNCLVARQHIESTM 300
Db	290 AAMLTGDDYLCDVWANBERISLQWLRLRIQYNSVMDIRDYNNTGKMISSAAQEHIESTM 349
Qy	301 TGRVYGRFRPSEPHFTDQNSFSYKISNEEGYRHICYFQIDKKDCDTFTKGTMEVIGEAL 360
Db	350 TGRVYGRFRPSEPHFTDQNSFSYKISNEEGYRHICYFQIDKKDCDTFTKGTMEVIGEAL 409
Qy	361 TSDLYXXISNEYKGMPPGRNLKYKQDLSYTKTCSELNBPERCQYVSFSKEAKYQ 420
Db	410 TSDLYXXISNEYKGMPPGRNLKYKQDLSYTKTCSELNBPERCQYVSFSKEAKYQ 469
Qy	421 RCGPGGLPYLTHSYNDKGIRVLEDNSALDKMLQVQMPSKCLDFILLNETKFWYQML 480
Db	470 RCGPGGLPYLTHSNDEELRVLEDNSALDKMLQVQMPSKCLDFILLNETKFWYQML 529
Qy	481 PPHFDKSKCKYKPLLDVYAGPCSQKADTYFRLNWATYLASTENIVASFDGRGSGYQGDKI 540
Db	530 PPHFDKSKCKYKPLLDVYAGPCSQKADTYFRLNWATYLASTENIVASFDGRGSGYQGDKI 589
Qy	541 MHAINRRLGTFEVDQTEAARQFSKMGFVDNKRIAIWGSYGGYNTSMVLGSGVPFKCG 600
Db	590 MEAVNRRLGTFEVDQTEAARQFSKMGFVDNKRIAIWGSYGGYNTSMVLGSGVPFKCG 649
Qy	601 IAVAPVSRWEYYSVTTERYGLPTPDNLHYRNTSMRAENFQKVEYLILHGTTADDN 660
Db	650 IAVAPVSRWEYYSVTTERYGLPTPDNLHYDYYKNTSMRAENFQKVEYLILHGTTADDN 709
Qy	661 VHFQSAQIISKALVYGVDFQAMWYTDDEDHGIASGPHQHIIYTHMSHFIKQCFSLP 716
Db	710 VHFAQSAQIISKALVYGVDFQAMWYTDDEDHGIASGPHQHIIYTHMSHFIKQCFSLP 765

RESULT 6	DEP4_PIG STANDARD; PRTR: 766 AA.
ID	DPP4_PIG
AC	P22411; Q866G2;
DT	01-AUG-1991 (Rel. 19, Created)
	01-AUG-1994 (Rel. 45, Last sequence update)
	25-OCT-2004 (Rel. 45, Last annotation update)
	13-SEP-2005 (Rel. 46, Last annotation update)
DE	Dipeptidyl peptidase 4 (EC 3.4.14.5) (Dipeptidyl peptidase IV) (Dipeptidyl peptidase 4 (EC 3.4.14.5)) (Contains: Dipeptidyl peptidase 4 (EC 3.4.14.5)) (Membrane activation antigen CD26) (Dipeptidyl peptidase IV membrane form) (Dipeptidyl peptidase IV membrane form); Dipeptidyl peptidase 4 soluble form (Dipeptidyl peptidase IV soluble form).
DE	Name=DEP4; Synonyms=CD26;
OS	Sub_ecrofa (Pig).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Suina; Suidae; Sub_ecrofa (Pig).
OC	Sub_ecrofa (Pig).
OX	NCBI_TaxID=9823;
RN	[1] TISSUE_Kidney;
RC	NUCLEOTIDE SEQUENCE, AND FUNCTION.
RP	PubMed=14719797; DOI=10.1515/BC-2003-172;
RX	Baer J., Weber A., Hoffmann T., Stork J., Wermann M., Wagner L., Aust S., Gerhardt B., Demuth H.-U.,
RA	"Characterization of human dipeptidyl peptidase IV expressed in <i>Pichia pastoris</i> . A structural and mechanistic comparison between the
RT	

recombinant human and the purified porcine enzyme.";

RT Biol. Chem. 384:1553-1563 (2003).
 RN [2] NUCLEOTIDE SEQUENCE OF 2-67.

RP TISSUE-Kidney;
 RC MEDLINE=9-1093209; PubMed=790569;
 RX Thomas P.-D., Qvist H., Marklund L., Andersson L., Sjostrom H.,
 RA Noren O.;
 RT "Assignment of the dipeptidylpeptidase IV (DPP4) gene to pig chromosome 15q11.";
 RL Mamm. Genome 4:604-607(1993).

RN [3] PROTEIN SEQUENCE OF 38-71.

RP TISSUE=Kidney;
 RC MEDLINE=91273813; PubMed=1675855;
 RX Seidl R., Mann K., Schaeffer W.;
 RT "N-terminal amino-acid sequence of pig kidney dipeptidyl peptidase IV solubilized by autolysis";
 RL Biol. Chem. Hoppe-Seyler 372:213-214 (1991).

RN [4] X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 39-766, N-GLYCOSYLATION SITE, AND HOMODIMERIZATION.
 RP PubMed=12690074; DOI=10.1073/pnas.0230620100;
 RA Engel M., Hoffmann T., Wagner L., Wermann M., Heiser U.,
 Kiefer-Bauer R., Huber R., Bode W., Demuth H.-U., Brandstetter H.;
 RT "The crystal structure of N-terminal dipeptidyl peptidase IV (CD26) reveals its functional regulation and enzymatic mechanism";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:5063-5068 (2003).
 CC -1- FUNCTION: Removes N-terminal dipeptides sequentially from polypeptides having unsubstituted N-termini provided that the penultimate residue is proline.
 CC -1- CATALYTIC ACTIVITY: Release of N-terminal dipeptide, Xaa-Yaa-[zaa-], from a polypeptide, preferentially when Yaa is Pro, provided CC -1- SUBUNIT: Homodimer.
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Also exists in a soluble form.
 CC -1- PTM: The soluble form (SDPP) derives from the membrane form (MDPP) by proteolytic processing.
 CC -1- SIMILARITY: Belongs to the peptidase S9B family. DPP4V subfamily.

This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

CC EMBL, AY198323; AAC043404.1; -; mRNA.
 DR EMBL; X7376; CA051717.1; -; mRNA.
 DR PIR; I47134; I47134;
 DR PDB; 1ORV; X-ray; A/B/C/D=39-766.
 DR InterPro; IPR003479; Ser. esters.
 DR InterPro; IPR00326; Peptidase_S9.
 DR PROSITE; PS00708; PRO-ENDOPEP_SER; 1.
 KW 3D-structure; Aminopeptidase; Direct protein sequencing; Glycoprotein; Hydrolase; Protease; Serine protease; Signal-anchor; Transmembrane.
 FT CHAIN 1 766 Dipeptidyl peptidase 4 membrane form.
 FT CHAIN 38 766 Dipeptidyl peptidase 4 soluble form.
 FT TOPO_DOM 1 6 Cytoplasmic (Potential).
 FT TRANSMEM 7 27 Signal-anchor for type II membrane protein (Potential).
 FT TOPO_DOM 28 766 Extracellular (Potential).
 FT ACT_SITE 630 630 Charge relay system (By similarity).
 FT ACT_SITE 708 708 Charge relay system (By similarity).
 FT ACT_SITE 740 740 Charge relay system (By similarity).
 FT CARBOHYD 85 85 N-linked (GlcNAc. .).

FT CARBOHYD 92 92 N-linked (GlcNAc. .). (By similarity).
 FT CARBOHYD 150 150 N-linked (GlcNAc. .). (Potential).
 FT CARBOHYD 179 179 N-linked (GlcNAc. .). (By similarity).
 FT CARBOHYD 219 219 N-linked (GlcNAc. .).
 FT CARBOHYD 229 229 N-linked (GlcNAc. .).
 FT CARBOHYD 279 279 N-linked (GlcNAc. .).
 FT CARBOHYD 321 321 N-linked (GlcNAc. .).
 FT CARBOHYD 393 393 N-linked (GlcNAc. .). (Potential).
 FT CARBOHYD 685 685 N-linked (GlcNAc. .).

FT DISULFID 385 394 Missing (in Ref. 2).
 FT DISULFID 444 447
 FT DISULFID 454 472
 FT DISULFID 649 762
 FT CONFLICT 32 32 88242 MW; 88000D>20BAE856D CRC64;
 SQ SEQUENCE 766 AA; 88242 MW; 88000D>20BAE856D CRC64;

Query Match 90.0%; Score 3489; DB 1; Length 766;
 Best Local Similarity 88.1%; Pred. No. 1.2e-229;
 Matches 45; Mismatches 40; Indels 0; Gaps 0;

Qy 1 NTYKLKLYSLRWLSDHELYKQENNLVFNAYEGNNSVPLENSTFDERGHSINDYNSISPD 60
 Db 51 STFRVKPFYTLQWISDHETLYKQENNLVFNAYEGNNSVLENSTFDERGHSINDYNSISPD 11.0
 Qy 61 GOFLLENNYVVKQRHHSYSTASYDYLANKRQLTTEERIPNNNTQWTVTSPVGHKLAYVWN 120
 Db 111 RQPLFENYVVKQRHHSYSTASYDYLANKRQLTTEERIPNNNTQWTVTSPVGHKLAYVWN 17.0
 Qy 121 DIYKLEPNLPSYRITWKGDLTYINGITDWVYEEEVSAWSAWWSPNGTFLAYAQFND 180
 Db 171 DIYKNEPNLSSQRTITWKGDKENVYINGITDWVYEEEVSAWSAWWSPNGTFLAYAQFND 23.0
 Qy 181 TEVPLIEFSYFSDESLOYPKTVPYPKAGAVNPVTKFEPVNTDSLSSVTNTSATQTAP 240
 Db 231 TEVPLIEFSYFSDESLQIPKTVRPIVYPKAGAENPTVKFEPVDTPLSPNASVTSYQIVPP 290
 Qy 241 ASMLIGDIYLCDVTTWATQERISLQWLRQIYQVSYMDICDYDESSGRMNCLVARYHIEST 300
 Db 291 ASVLGDIAYLCGTWVTEBRIQWIRIAQNEYTIDICDYDESGRMRVSSVQHQIBIST 350
 Qy 301 TGWYGRFRPSEPHFTLDGNSFYKIIISNEGYRHICYQFLDKDCTPITKGTVWVIGIRAL 360
 Db 351 TGWYGRFRPAPFITSDDNSFYKIIISNEGYRHICFHQTDSKNTCPTIKGAVEWVIGIRAL 410
 Qy 361 TSDLYYIISNEYKGMPGGRNLKYKQLSDYTKVTCLSCELNPERCQYSVSFSKEAKYQYL 420
 Db 411 TSDWLYTISNEHGMPGRNLTRIQLNXYTRVTCLSCELNPERCQYSASFSNKARYQYL 47.0
 Qy 421 RCGSPGLPLYTHLSYVNDKGLRLVEDNSALDKMLQNYQMPSKLQLDFITLNPKFWYQML 480
 Db 471 RCGPGLPLYTHLSSSDKELRTEEDNSALDKMLQDVQMPSKLQLDVNLHGTPFWYQML 53.0
 Qy 481 PPFDKSKKTPLLDYYAGPCSKADYFPLNATYLASTENIVASPDGRGSGGGYQGDKI 54.0
 Db 531 PPFDKSKKTPLLIENYAGPCSKVDTFPLRSLWATYLASTENIVASPDGRGSGGVFCGG 60.0
 Qy 541 MHAINRRGTFEVDOTEAAAROFSKMGFVDNTRIAIANGWSYGGYVTSMVLSGSGLGVFKCG 660
 Db 591 MHAINRRGTFEVDQEAATROFSKMGFVDNTRIAIANGWSYGGYVTSMVLSGSGLGVFKCG 65.0
 Qy 601 IAVAPVSRWEYDSVYTERYMLPPTEDNLHYRNSTMSRAENFKQVBYLLIHTADDN 71.0
 Db 651 IAVAPVSRWEYDSVYTERYMLPPTENLDYRNSTMSRAENFKQVBYLLIHTADDN 71.0
 Qy 661 VHQQSAQISKALVYDGYDFOAMWYTDHDGLASSTAHOHITYTHMSHPIKOCPSLP 716
 Db 711 VHQQSAQISKALVYDGYDFOAMWYTDHDGLASSTAHOHITYTHMSHPIKOCPSLP 766

RESULT 7
 DPP4 RAT
 ID DPP4 RAT
 STANDARD: PRT: 767 AA.
 AC P14740;

DT	01-APR-1990 (Rel. 14, Created)	CC	Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided
DT	01-FEB-2005 (Rel. 46, Last sequence update)	CC	Zaa, is neither pro nor hydroxyproline.
DT	13-SEP-2005 (Rel. 48, Last annotation update)	CC	-1- SUBUNIT: Homodimer.
DE	Dipeptidyl peptidase 4 (BC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP IV)	CC	-1- SUBCELLULAR LOCATION: Type II membrane protein. Also exists in a soluble form.
DE	T-cell activation antigen CD26	CC	-1- TISSUE SPECIFICITY: Expressed in bile ducts and other epithelial brush borders (small intestine, kidney, colon, pancreatic duct); acinar structures in salivary glands; endothelial structures and T cell areas in thymus, spleen and lymph node.
DE	canalicular domain-specific membrane glycoprotein [Contains: Dipeptidyl peptidase 4 membrane form (Dipeptidyl peptidase IV form); Dipeptidyl peptidase 4 soluble form (Dipeptidyl peptidase IV soluble form); Dipeptidyl peptidase 4 60 kDa soluble form (Dipeptidyl peptidase IV 60 kDa soluble form)].	CC	-1- PTM: The soluble form (SPP) derives from the membrane form (MDPP) by proteolytic processing.
OS	Name=Dpp4; Synonyms=CD26;	CC	-1- SIMILARITY: Belongs to the DPPIV subfamily.
OC	Eukaryota; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Rattus.	CC	This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.
RN	Rattus norvegicus (Rat); NCBI_TAXID=10116;	CC	This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.
RX	NUCLEOTIDE SEQUENCE, AND PARTIAL PROTEIN SEQUENCE. MEDLINE=9123496; PubMed=2533382;	CC	DR EMBL; AAA41096.1; -; mRNA.
RA	Ogata S., Misumi Y., Ikebara Y.; "Primary structure of rat liver dipeptidyl peptidase IV deduced from its cDNA and identification of the NH2-terminal signal sequence as the membrane-anchoring domain.", J. Biol. Chem. 264:3596-3601 (1989).	DR EMBL; J04521; AAA41272.1; -; mRNA.	
RT	[1]	DR PIR; A39914; A39914.	
RT	RX	DR HSSP; P27487; 1PFQ.	
RT	RN	DR MEROPS; S09.003; -.	
RT	RN	DR Ensembl; ENSRNOG00000030763; Rattus norvegicus.	
RP	NUCLEOTIDE SEQUENCE. MEDLINE=8806516; PubMed=3479775;	DR RGD; 2515; DPP4.	
RX	Hong W., Doyle D.; "cDNA cloning for a bile canaliculus domain-specific membrane glycoprotein of rat hepatocyte.", Proc. Natl. Acad. Sci. U.S.A. 84:7962-7966 (1987).	DR GO; 0005625; C:membrane fraction; IDA.	
RA	RN	DR GO; 0005625; C:soluble fraction IDA.	
RT	"Membrane orientation of rat gp110 as studied by in vitro translocation.", J. Biol. Chem. 263:16892-16898 (1988).	DR CO; 0004224; F:dipeptidyl-peptidase IV activity; TAS.	
RT	RN	DR GO; 0006508; P:peptidolysis and peptidolysis; TAS.	
RT	RN	DR InterPro; IPR002471; Pept_S9_AS.	
RP	NUCLEOTIDE SEQUENCE OF 1-40. MEDLINE=9003185; PubMed=3182821;	DR InterPro; IPR001375; Pept_Iidae_S9.	
RX	Hong W.-J., Doyle D.; "Identification of the bile canalicular cell surface molecule GP110 as the ectopeptidase dipeptidyl peptidase IV: an analysis by tissue distribution, purification and N-terminal amino acid sequence.", Hepatology 11:534-544 (1990).	DR InterPro; IPR00469; Peptidase_S9.	
RA	RN	DR InterPro; IPR000379; Ser esterase.	
RT	RN	DR Pfam; PF00930; DPPIV_N; 1.	
RT	RN	DR PROSITE; PS00708; PRO-ENDOPEP_SER; Peptidae_S9.	
RT	RN	DR KW Aniopeptidase; Direct protein sequencing; Glycoprotein; Hydrolase; Protease; Serine protease; Signal-anchor; Transmembrane.	
RP	PROTEIN SEQUENCE OF 28-58, AND TISSUE SPECIFICITY. MEDLINE=20228896; PubMed=11970322;	DR Dipeptidyl peptidase 4 membrane form.	
RX	McCaughan G.W., Wickson J.E., Creswick P.F.; Gorrell M.D.; "Identification of the active site residues in dipeptidyl peptidase IV.", J. Biol. Chem. 274:973-975 (1993).	DR Dipeptidyl peptidase 4 soluble form.	
RA	RN	DR Dipeptidyl peptidase 4 60 kDa soluble form.	
RT	RN	DR Cytoplasmic (Potential). Signal-anchor for type II membrane protein (Potential).	
RT	RN	DR Extracellular (Potential). Charge relay system (By similarity).	
RP	PROTEIN SEQUENCE OF 281-302, AND MUTAGENESIS OF GLY-629; TRP-630; SER-631; TRY-632 AND GLY-633. MEDLINE=94128329; PubMed=79052721;	DR Charge relay system (By similarity).	
RX	Iwaki-Egawa S., Watanabe Y., Fujimoto Y.; "N-terminal amino acid sequence of the 60-kDa protein of rat kidney dipeptidyl peptidase IV.", J. Biol. Chem. 31:2582-2587 (1992).	DR N-linked (GLCNAC, .) (By similarity).	
RA	RN	DR N-linked (GLCNAC, .) (By similarity).	
RT	RN	DR N-linked (GLCNAC, .) (By similarity).	
RT	RN	DR N-linked (GLCNAC, .) (By similarity).	
RT	RN	DR N-linked (GLCNAC, .) (By similarity).	
RT	RN	DR N-linked (GLCNAC, .) (By similarity).	
RP	PROTEIN SEQUENCE OF 624-648. MEDLINE=2190188; PubMed=11347701;	DR N-linked (GLCNAC, .) (By similarity).	
RX	Ogata S., Misumi Y., Tsuji E., Takami N., Oda K., Ikebara Y.; "Identification of the active site residues in dipeptidyl peptidase IV by affinity labeling and site-directed mutagenesis.", Biochemistry 31:2582-2587 (1992).	DR N-linked (GLCNAC, .) (By similarity).	
RA	RN	DR By similarity.	
RT	RN	DR By similarity.	
RT	RN	DR By similarity.	
RT	RN	DR By similarity.	
RP	SIGNAL-ANCHOR. MEDLINE=90338089; PubMed=11974258; DOI=10.1083/jcb.111.2.323;	DR G->A: Reduced activity.	
RX	Hong W., Doyle D.; "Molecular dissection of the NH2-terminal signal-anchor sequence of rat dipeptidyl peptidase IV.", J. Cell Biol. 111:323-328 (1990).	DR G->R: Reduced activity.	
RA	RN	DR W->E: No effect on activity.	
RT	RN	DR S->A: Reduced activity.	
RT	RN	DR Y->F: No effect on activity.	
RT	RN	DR Y->G: Reduced activity.	
RP	"FUNCTION: Removes N-terminal dipeptides sequentially from polypeptides having unsubstituted N-termini provided that the penultimate residue is proline.	DR Y->L: Reduced activity.	
CC	"CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa." -	DR G->A: Reduced activity.	
CC	"CATALYTIC ACTIVITY: Removal of an N-terminal dipeptide, Xaa-Yaa." -	DR G->S: Reduced activity.	
CC	"CATALYTIC ACTIVITY: Removal of an N-terminal dipeptide, Xaa-Yaa." -	DR R -> A (in Ref. 1).	
RN	CONFLICT	DR 38	

FT	CONFLICT	54	54	Missing (in Ref. 4).
FT	CONFLICT	183	183	I -> T (in Ref. 2).
FT	CONFLICT	332	332	T -> N (in Ref. 2).
FT	CONFLICT	352	352	C -> V (in Ref. 2).
FT	CONFLICT	394	394	V -> D (in Ref. 2).
FT	CONFLICT	562	562	L -> F (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
SQ	SEQUENCE	767	AA:	88089 MW: ED94717RF1B3E40 CRC64;
Query Match	Score: 3354.5;	DB: 1;	Length: 767;	
Best Local Similarity	84.7%;	Pred. No.	1.9e-220;	
Matches 608; Conservative	52;	Mismatches	55;	Indels 3; Gaps 1;
QY	1	NTYRLKYSLRWISDHELYLQKENNLVENNAEYGNSSVPELENSTEDFGHSINDYSISPD 60		(Dipeptidyl peptidase IV soluble form)].
Db	49	NTFRVKSYSLRWISDSELYLQKENNLVNAEHGNSSIFLENSTEIFGDSISDVSPPD 108		Name=Dpp1; Synonyms=Cd26;
QY	61	GQFILELINTVKQWHSYTASYDIDYLKLQLITEBRIPNNTQWTVSPVGHKLAYWN 120		Mus musculus (Mouse)
Db	109	RLFVLILEVNYYKQWHSYTASYDIDYLKLQLITEEKIPNNTQWTVSPQEGHKLAYWN 168		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Burchontoglires; Glires; Rodentia; Sciurognathia; Murioidea; Muridae; Murinae; Mus.
QY	121	DIVYKIEPNPFSYRTRWPEKDTIINGITWVYEEFVSYASALWWSPICTFLAYAQFD 180		NCBI TaxID=10090;
Db	169	DIVYKIEPNPFSHRTSGENVINGINWVYEEFVSYASALWWSPICTFLAYAQFD 228		NCBI TaxID=1;
QY	181	TEVPLIEFSYPSDESQYKPTKVRPVPKAGAVANPVIKFKEVNTDSLSSVTNATSIQTAP 240		DE STRAIN=Swiss; TISSUE=Thymus;
Db	229	TGVPLIEFSYPSDESQYKPTKVRPVPKAGAVANPVIKFKEVNTDSLSSTTIMQITAP 288		MEDLINE=0212288; PubMed=1370813;
QY	241	ASMLIGDHYLCDVTAQERISLQWLRIONYSWMDICDYEDESSRMNLVAROHIENST 300		RA Marguet D.A., Bernard A.-M., Vivier I., Darmoul D., Naquet P., Pierres M;
Db	289	ASVTTGHDYLCDVAMVEDRISLQWLRIONYSWMAICDYDKTTLVNNCPTQHIEISA 348		RA "cd26 cloning for mouse thymocyte-activating molecule. A multifunctional ecto-dipeptidyl peptidase IV (cd26) included in a subgroup of serine proteases."
QY	301	TGWLWGRFRPSEPHFTLDGNSFYKITSNEEGYRHICYFQDJKD--CPTITKETWEGI 357		RA Structure of the mouse dipeptidyl peptidase IV (cd26) gene.";
Db	349	TGWLWGRFRPSEPHFTSGSFYKIVSDQGKHYCQFKQRDRKPQVCFTIKGAWEVSI 408		RA J. Biol. Chem. 267:2200-2208 (1992).
QY	358	EALISDLYLVIISNEYKGMPCGRNLKYKIQSLDPTVTCSCELNPERCQYSVSFSKEY 417		RA [2]
Db	409	EALTSDDLYLVIISNEYKEMPGCRNLKYKIQSLDPTVTCSCELNPERCQYSVSFSKEY 468		RA RP SEQUENCE REVISION.
QY	418	YOLRSGGPGLPYLTHSVDNGKGRVLEDNALDKMLQNQMPSKRKLDFILNETKPYQ 477		RA RA Marguet D.A.; Bradt D., Erusic V., Chothia C., Corbani L.E., Cousins S., Dala E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S., Gaasterland T., Garibaldi M., Giessi C., Godzik A., Gough J., Groommond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D., Karai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L., Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A., Maglott D.R., Maitais L., Marchonni L., McKenney L., Miki H., Nasashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G., Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S., Rasaki T., Reed J.C., Reid J., Ringwald M., Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K., Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M., Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C., Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang J., Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N., Hirono-Kishikawa T., Kondo H., Nakamura M., Sakazume N., Sato K., Shiraki T., Waki K., Kawai J., Aizawa T., Fukuda S., Hara A., Hashizume W., Imori K., Ishii Y., Itoh M., Kagawa I., Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinsagawa A., Yuzunishi A., Yoshino M., Landre R., Waterston R., Rogers J., Birney E., Hayashizaki Y.; RT "Analysis of the mouse transcriptome based on functional annotation of the mouse genome."
Db	469	YQLGRGPGLPYLTHRSTDQKELRVLEDNALDKMLQDVQMPSKKLDFIVLNTRFWTQ 528		RA RT 60-70 full-length cDNAs."/>
QY	478	MILPHPDFSKSKYKPYLLDYYAGPCSQADTYFLRNLWATLASTENITIASFDGRGSQIG 537		RA RN [5]
Db	529	MILPHPDFSKSKYKPYLLDYYAGPCSQADTYFLRNLWATLASTENITIASFDGRGSQIG 588		RA RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
QY	538	DKIMHAINRRLGTPEVEDQTEAARQFSKMGFPUDKNKRIAWGMWSGGYVTSMVLGSGGVF 597		RA RX MEDLINE=22308257; PubMed=101073/pnas.242603899;
Db	589	DKIMHAINRRLGTPEVEDQTEAARQFSKMGFPUDKNKRIAWGMWSGGYVTSMVLGSGGVF 648		RA RA Klausberg R.L., Peingold B.A., Grouse L.H., Derge J.G., Klausberg R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buerow K.H., Schaefer C.P., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Staapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheerz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loqueilano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosek S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
QY	598	KCGIAVAPSRWEYDSDVTERYGLPFPEDNLHYRNTVMSRAENPKQVEYLHHTA 657		RA
Db	649	KCGIAVAPSRWEYDSDVTERYGLPFPEDNLHYRNTVMSRAENPKQVEYLHHTA 708		RA
QY	658	DDNVHFQSOAQSKALVDPGVDFOAMWYTEDDGHIASSTAHQHLYTASHFETIKOCFSU 715		RA
Db	709	DDNVHFQSOAQSKALVDAVGVDQAMWYTEDDGHIASSTAHQHLYTASHFLOOCFSU 766		RA
RESULT 8				
DPP4_MOUSE	STANDARD;	PRT;	760 AA.	
ID_P28843_MOUSE				
AC_P28843	1992 (Rel. 24, Created)			
DT_01-FEB-1992	(Rel. 33, Last sequence update)			
DT_01-FEB-1996	(Rel. 48, Last annotation update)			
DT_13-SEP-2005	(Rel. 48, Last annotation update)			
DE_Dipeptidyl peptidase 4 (EC 3.4.14.5) (Dipeptidyl peptidase IV) (T-cell activation antigen CD26) (Thymocyte-activating molecule)				
DE_(THAM) (Contains: Dipeptidyl Peptidase 4 membrane form (Dipeptidyl peptidase IV membrane form); Dipeptidyl peptidase 4 soluble form				

Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Murry D.M., Sodegeren E.J., Lu X., Gibbs R.A., Fahey J., Heiton B., Ketteman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Rodiguez S., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J.W., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E., Schnarch A., Schein J.B., Jones S.J.M., Marra M.A.; "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."; *Proc. Natl. Acad. Sci. U.S.A.* 99:16899-16903 (2002). [6]

PROTEIN SEQUENCE OF 1-20.
MEDLINE:91302787; Pubmed:1712807;
RA Vivier L., Marguet D.A., Naquet P., Bonicel J., Black D., Li C.X.-., Bernard A.-M., Gorvel J.-P., Pierres M.; "Evidence that thymocyte-activating molecule is mouse CD26 (dipeptidyl peptidase IV)."; *J. Immunol.* 147:447-454 (1991).
CC -I- FUNCTION: Removes N-terminal dipeptides sequentially from polypeptides having unsubstituted N-termini provided that the penultimate residue is proline.
CC -I- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-[-zaa], from a polypeptide, preferentially when Yaa is Pro, provided Zaa is neither Pro nor hydroxyproline.
CC -I- SUBUNIT: Homodimer.
CC -I- SUBCELLULAR LOCATION: Type II membrane protein. Also exists in a soluble form.
CC -I- PTM: The soluble form (SDPP) derives from the membrane form (MDPP) by proteolytic processing.
CC -I- SIMILARITY: Belongs to the peptidase S9B family. DPPV subfamily.

This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

DR EMBL; X59384; CAA11274.1; "-"; mRNA.
DR EMBL; U12620; AAA82213.1; "-"; Genomic_DNA.
DR EMBL; U12599; AAA82213.1; "-"; Genomic_DNA.
DR EMBL; U12600; AAA82213.1; "-"; Genomic_DNA.
DR EMBL; U12601; AAA82213.1; "-"; Genomic_DNA.
DR EMBL; U12602; AAA82213.1; "-"; Genomic_DNA.
DR EMBL; U12603; AAA82213.1; "-"; Genomic_DNA.
DR EMBL; U12604; AAA82213.1; "-"; Genomic_DNA.
DR EMBL; U12605; AAA82213.1; "-"; Genomic_DNA.
DR EMBL; U12607; AAA82213.1; "-"; Genomic_DNA.
DR EMBL; U12608; AAA82213.1; "-"; Genomic_DNA.
DR EMBL; U12609; AAA82213.1; "-"; Genomic_DNA.
DR EMBL; U12610; AAA82213.1; "-"; Genomic_DNA.
DR EMBL; U12611; AAA82213.1; "-"; Genomic_DNA.
DR EMBL; U12612; AAA82213.1; "-"; Genomic_DNA.
DR EMBL; U12613; AAA82213.1; "-"; Genomic_DNA.
DR EMBL; U12614; AAA82213.1; "-"; Genomic_DNA.
DR EMBL; U12615; AAA82213.1; "-"; Genomic_DNA.
DR EMBL; U12616; AAA82213.1; "-"; Genomic_DNA.
DR EMBL; U12617; AAA82213.1; "-"; Genomic_DNA.
DR EMBL; U12618; AAA82213.1; "-"; Genomic_DNA.
DR EMBL; U12619; AAA82213.1; "-"; Genomic_DNA.
DR EMBL; AK085370; BAC39434.1; "-"; mRNA.
DR HSSP; P27487; 1PFQ.
DR SNR; P28843; 37-759.
DR MEROPS; S09_003; "-".
DR Ensembl; ENSMUSG0000035000; *Mus musculus*.
DR GO; GO:0016021; C: integral to membrane; TAS.
DR GO; GO:0046581; C:intercellular canalicular; IBA.
DR InterPro; IPR002471; Pept_S9_AS.
DR InterPro; IPR001375; PeptDage_S9.
DR InterPro; IPR002469; Peptidase_S9B.

DR InterPro; IPR000379; Ser_estr.
DR Pfam; PF00530; DPPIV_N_1.
DR PROSITE; PS00705; PRO_ENDOPEP SER_1.
KW Aminopeptidase; Direct protein sequencing; Glycoprotein; Hydrolase; Proteinase; Serine protease; Signal-anchor; Transmembrane.
FT CHAIN 1 760
FT CHAIN 37 760
FT TOPO_DOM 1 6
FT TRANSMEM 7 28
FT TOPO_DOM 295 466
FT COMPBAS 624 624
FT ACT_SITE 702 702
FT ACT_SITE 734 734
FT CARBOHYD 83 83
FT CARBOHYD 90 90
FT CARBOHYD 213 213
FT CARBOHYD 223 223
FT CARBOHYD 315 315
FT CARBOHYD 328 328
FT CARBOHYD 514 514
FT CARBOHYD 679 679
FT DISULFID 379 388
FT DISULFID 438 441
FT DISULFID 448 466
FT DISULFID 643 756
SQ 760 AA; 87437 MW; ASF64B465A3DF8 CRC64;

Query Match Score 86.1%; DB 1; Length 760;
Best Local Similarity 84.9%; Prede. No. 1-9e+219; Mismatches 49; Indels 4; Gaps 1;
Matches 607; Conservative 55; N-linked 1; O-linked 1; Gap 1;

Qy 1 NTYRKLYSLRWNISHDHELYKQENNLVFNAYEGNSSELENSTDEFGHSINDYSTSPD 60
Db 49 STFRVTSYSLWNVSDFSYLYKQENTNLLNAEHGNSIFLENSTFESFGY --- -HSVSPD 104

Qy 61 GQPILLENNYKQWRISRTASYDIDLNKRQLTIEPIPNNTQWTWSPVGHKLAYWN 120
Db 105 RLFLVLENNYKQWRISRTASYNDVNKRQLTIEEKIPNNTQWTWSPSFGHKLAYWN 164

Qy 121 DIVXKTEPNLPSYRITWGRKDLYINGITDWVYEEVFSAWSNPGTFLAYAFND 180
Db 165 DIVXKTEPHLPSRITSGEENVITGIDWVYEEFEEFGASALWSPNTFLAYAFND 224

Qy 181 TBYPLIEYSFSDESLOQPKTVTPYKPGAVNPYTKFFVNTDSLSVTNTSQTAP 240
Db 225 TGYPLIEYSFSDESLOQPKTVWIPYKPGAVNPYTKFFVNTDISLSSSSAAPIQPAP 284

Qy 241 ASMLIGPHYLCDVTWATQERISLWQRRIQNSVMDICDYDESSGRWNLVARYHEMST 300
Db 285 ASVARGDHYLCDVWATERISLQWRRRIQNSVMAICDYDKINLTNCPSQQHVEMST 344

Qy 301 TGWGRFRPSEPHFTLDGNSFSYK1ISNEEGYRHYCIFYFQIDRKDCTFITKGTMWVEIGEAL 360
Db 345 TGWGRFRPAEHPFGSSFTK1ISDGKPKDDCTFTIKGAEVISTEAL 404

Qy 361 TSDLYYYISNEYFKGMPGGRNLKYKQIOLSDTYTKVTCBLNPERCQYVSFSKEAKYQL 420
Db 405 TSDLYYYISNQYKMPGGRNLKYKQIOLTDHTNVKCLSDLNPERCQYAVSFKEAKYQL 464

Qy 421 RCGSGPLPYTHSYNDKGGRVLEDNSALDKMQLQYOMPSKLDFTLNETKWTOMIL 480
Db 465 GCWGPGLPYTLHRSTPHKELRVLESALORLMDQYOMPSKLDFTLNETRFYQML 524

Qy 481 PPHFDKSRSKPKYQPLDLYAGPCSKQADYVFRNLTAYLASTENIVASFDGROSSGYQGDKI 540
Db 525 PHFDKSRSKPKYQPLDLYAGPCSKQADYVFRNLTAYLASTENIVASFDGROSSGYQGDKI 584

Qy 541 MHAINRRLGTFEVEDQEAAOFSKNGFVDNKRIAIWGWSYGGYTSMVLSGSGVFKCG 600

Qy	63 FILILEYNTVKQWRHSYTASYDIYDLNKROLITERIPNINTOWWSPVGHKLAYVNNDI	122	DR GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.
Db	107 FALIQYSTEKWRHSYTASYHIDLNNTRKITENPLPNIQYIWSWPVGHKLAYVNIN	166	DR GO; GO:000508; P:proteinase and peptidylsisis; IEA.
Qy	123 YVKIEPNLPSYRITWTGKEDIYNGITDWWYEEFVSSAYSALMWSPNCTELAYAQFNDE	182	DR InterPro; IPR002471; Pept_S9_AS.
Qy	167 YVKATPNSAPVQTENGKAQNKLINGLAWDVYEEBEMGTHSALWSPNCFALAEINDE	226	DR InterPro; IPR001375; Peptidase_S9_B.
Db	183 VPLIEYNTSDESLOQPTTRVVPKAGAVNPFTKFFTVNTSLSSVNTNATIQLTAPAS	242	DR InterPro; IPR000379; Ser_estrs.
Qy	227 VPWBEPSPYSEDTIQLQPCKTKIKIPKPKAGAINPTRLFLVDI---SLSPKNSEIVASS	282	DR Pfam; PF000326; Pept_Dase_S9_1.
Qy	243 MLIGDHYLCLDVTWATERISLOWLRRIONYSYNDICDYEDESSGRWNCLVAROHIEEMSTG	302	DR Pfam; PS00708; PRO_Endopeptidase_ER; 1.
Db	283 IISSEDHYSLSVWITWTDERLVCWMARRINSFVLTICDY---SGAWHCKPERHELEESTG	339	KW Hydrodase; Hypothetical protein; Transmembrane.
Qy	303 WVGFRPSEPHTLDGNSFYKILISNEEGSYRHICYFOIDKDCTFITKGTMWIGEIALTS	362	SQ SEQUENCE 759 AA; 86343 MW; B9A0C38CCB8644AC CRC64;
Db	340 WVGFRQSPSEPYFSDKISYRISDFKSYRISDFKSYRISDFKSYRISDFKSYRISAVTN	398	DR Best Local Similarity 64.6%; Pred. No. 1.5e-160; Mismatches 147; Indels 16; Gaps 9;
Qy	363 DLYIXISNEYKGPGGRHLYKIQSLDTPKTCUSCELMPPERCOYKYSFSESKARYTQRC	422	DR Matches 466; Conservative 92; Mismatches 146; Indels 16; Gaps 9;
Db	399 NSLYFISNEFEGRGGRHLYKVDLKDNUKKICITCNCSCEEAQCQFSVFSSTDSSRYKLN	458	DR 50 NTYRLKLYSLRMRNISDHELYKOEN-NILVFNAEYGNSSVYLENSTFDEFGHSINDYDSTP 59
Qy	423 SGPSPLTTLHSSYNDKGURVLDNBDSADLDRMLNQVMMPSKQLDFTILNBTKFWYQMLLPP	482	DR 50 NDYVYKTHNLQWTSQGNOYLHETSNGNGLRFLDAETGTSVVLLNTISTI--HEATAILS 107
Db	459 YGDPDPYFTLQNSSTDKAITKLEEDNNKRNKVNKEIQMFCRLLSNVLTQITWQMILLP	518	DR 60 DQFQILLEYNTKQWRHSYTASYDIDLNKRQLOLITERIPNINTOWWSPVGHKLAYVNIN 119
Qy	483 HFDKSKKCPKLDDVYAGPCGSQRADTFVNLNWATLASTENILVASFDGRGSGYQGDKLMH	542	DR 108 DQRFAUQQYKVEKLWRSYTAHYIDNTTSILDDALLPNDTQYIWSWPVGHKLAYVNIN 167
Db	519 NFDSSKCPKLDDVYAGPCGSQKADAIFINWSTYLASEBGIIVASFDGRGSGFQGDCLH	578	DR 120 NDIVYKLEBPNLPSYRITWTGKEDIITYNGITDWWYEEFVSSAYSALMWSPNCTELAYQFN 179
Qy	543 AINBRIGTBEVEDOIAEABOSRSMGFGDNKRINWGGYGTTSMSLJGSSGCVFKCJA	602	DR 168 NTYIKASPTARPVQITSNGEENKINPQDNPVYEEBMFGSISALATSPGNFVAIAFPN 227
Db	579 AIYRRLGTYEVEDQISAAKLFSEMSFVKDRIAWGNYYGGYTSWMLGAGSDVFKCGIA	638	DR 180 DTEVPPLIEYSFSDESLQYQPKTIVRVPYPKAGAVNPFTKFFVNTDSLSSVTNATSIOT 239
Qy	603 VAPSRWYEDSYTERWGLPPEDNIDHYRNSTMSRAENPKQVEWLLHGTADDNVH	662	DR 228 DTEVPVBYSPFSDEDTIQLQPKTIRIPKAGAKNPKTVRKFVLD--.QMLPDNPSTEISP 284
Db	639 VAPSRWYEDSYTERWGLPKENDNFYENSTVPAKRFPTDYLINGTADDNVH	698	DR 240 PASMIGDHYLCLDVTWATERISLOWLRRIONYSYNDICDYEDESSGRWNCLVARQH--I 296
Qy	663 FOOSAQIISKALVDPVGVDFOAMMYTDEDGIASSTAHQHIIYTHMSHEFKQCFSLP	716	DR 285 PAEKSCDHYLSVVTWYDTERCLQWLRRIONYSYVLTCDFPSATGNWTC--HRKDNFWK 342
Db	699 FOQAQIISKALVDAEVDFQAMWYTDQDGIGGH-AHSIIYQHMSHFMRQCFKLP	751	DR 297 EMSTGHWGRFRPSEPHFTLDNSYKLIISNEBGYRHICYFQIDKQDCTFITKGTVWEIG 356
Qy	RESULT 1.1 QZ181 CHICK PRELIMINARY; PRT: 759 AA.	751	DR 343 KVIQLAGW -QISAIACFLPAPBNTYYKVSFVSYTGKHYHINGTEQAVPV-TIEGKSFV 399
QZ181	QZ181_ CHICK PRELIMINARY;	759	DR 357 TEALTSQDLYTTSNEYKGMPCGGRNLVYKQIOLSDYKTYK-CLSCELNPRCOYYSVSFSKEA 415
AC	AC QZ181_		DR 400 IAATVKFLYYISNQNCMPGPNRLYKMLLBESPSTQCVSDLNQRCQYISASFSDA 459
DT	DT 25-OCT-2004 (TREMBrel. 28. Created)		DR 416 KXYQLRQSGGPGLIPLYIHTHSSYNDKGRLPTEEDNSALDMLQVOMPSKQKLDEFILNEKF 475
DT	DT 25-OCT-2004 (TREMBrel. 28. Last sequence update)		DR 460 QXQQLNC1GPGLPMSTLHRSSDQVRLYLENTELENSLKDQMPSKQKLGSITVGGYNLW 519
DE	DE ORFNames=RCJMB04_29g21;		DR 476 YOMILPHPFDKSKKXPPLDLYVAGPCSQKADTFVNLNWATLASTENILVASFDGRGSGY 535
OS	OS Gallus gallus (Chicken);		DR 520 YOMILPHFLDSSSKYKPFLLTYYAGPSQKDVH弗VLAETQIVAVSFDFGRGSGY 579
OC	OC Archosauaria; Aves; Chordata; Craniata; Galliformes; Phasianidae; Phasianinae;		DR 536 QGDKIMAHAINRRLGTFFEVDQLEAEOFSKMGCFVNDRKRIAIWGSICGGYVTSMLVGSGS 595
OC	OC Gallus gallus.		DR 580 QGDIMAHAINRRLGTFEVDQISARTFSEMSFVDKDRIAIWGSICGGYVTSMLVGSGS 639
RN	RN NCBI_TaxID=9031;		DR 596 VPKCGIAYAVPYSRWYEDSYTERWGLPPEDNIDHYRNSTMSRAENPKQVEWLLHJHG 655
RP	RP NUCLEOTIDE SEQUENCE.		DR 640 VPKCGIAYAVPYSRWYEDSYTERWGLPPEDNIDHYRNSTMSRAENPKQVEWLLHJHG 699
RC	RC STRAIN=CB; TISSUE=Bursa;		DR 656 TADDNYHFOQSOQISKALVDPVGVDFOAMWYTDQDGIASSTAHQHIIYTHMSHP1KQCFSL 715
RA	RA Caldwell R.B., Kierzek A.M., Arakawa H., Bezzubov Y., Zaim J.,		DR 700 TADDNYHFOQSAQISKALVDAEVDFQAMWYTDQDGIGGH-AHSIIYQHMSHFMRQCFKLP
RA	RA Friedler P., Kutter S., Blagodatski A., Kostovska D., Kotter M.,		DR 716 P 716
RA	RA Plachy J., Carnici P., Hayashizaki Y., Buerstedde J.M.;		DR 759 P 759
RT	RT "Full-length cDNA from chicken bursal lymphocytes to facilitate gene function analysis."		DR RESULT 12
RR	RR Genome Biol. 6:R6-R6 (2005).		DR 0641D6_XENLA PRELIMINARY; PRT; 737 AA.
DR	DR EMBL: AJ720303; OAG32562.1; mRNA.		DR ID Q641D6_XENLA PRELIMINARY; PRT; 737 AA.
DR	DR GO:0004177; C:membrane IEA.		DR AC Q641D6_XENLA PRELIMINARY; PRT; 737 AA.
DR	DR GO:00004274; F:dipeptidyl-peptidase IV activity; IEA.		
DR	DR GO:0016787; F:hydrolase activity; IEA.		

DT 25-OCT-2004 (T-EMBLrel. 28, Created)
 DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)
 DE LOC397888 protein.
 CN Name=LOC397888;
 CS Xenopus laevis (African clawed frog)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
 OC Xenopodinae; Xenopus; Xenopus;
 OX NCBI_TaxID=8355;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Kidney;
 RX MEDLINE=23341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Straubberg R.L., Wagner L., Pontius J., Clifton S.W.,
 Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 initiative.";
 RL Dev. Dyn. 225:384-391 (2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Kidney;
 RX MEDLINE=2238857; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Straubberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 Klausner R.D., Collins F.S., Wagner L., Shemesh C.M., Schuler G.D.,
 Altschul S.F., Zeeberg B., Bluetow K.H., Schaefer C.F., Bhat N.K.,
 Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 Raha S.S., Logueillo N.A., Peters G.J., Abramson R.D., Mallahy S.J.,
 Bosak S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 Fahy J., Hellton E., Kertmann M., Madan A., Rodrigues S., Sanchez A.,
 Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 Rodriguez A.C., Grimwood J.W., Schmutz J., Myers R.M.,
 Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smilus D.E.,
 Schnierch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Kidney;
 RA Klein S., Gerhard D.S.; Submitted (Sep 2004) to the EMBL/GenBank/DDBJ databases.
 EMBL: BC082401; AAH82401.1; -; mRNA.
 DR GO:0016020; C:membrane; IEA.
 DR GO:0004177; C:faminopeptidase activity; IEA.
 GO: GO:0004224; F:dipeptidyl-peptidase IV activity; IEA.
 DR GO: GO:0016287; F:hydrolase activity; IEA.
 DR GO: GO:0004287; F:prolyl oligopeptidase activity; IEA.
 DR GO: GO:0006508; P:protein processing and peptidolysis; IEA.
 DR InterPro; IPR002471; Pept-S9 AS.
 DR InterPro; IPR001375; Peptidase_S9.
 DR InterPro; IPR002469; Peptidase_SS5.
 DR Pfam; PF00930; DPPIV_N_1.
 DR PROSITE; PS00700; PRO-ENDOPBP_SER; 1.
 KW Hydrolase; Transmembrane.
 SQ SEQUENCE 737 AA; 84184 MW; 1E08CP94EDFBCC4B CRC64;

Query Match 59.6%; Score 2309; DB 2; Length 737;
 Best Local Similarity 59.1%; Pred. No. 5e-149; Mismatches 164; Indels 20; Gaps 9;

3 YRLKLYSPRWSIDSEYLYQENNIVLVEAYGNSSVFLENSTPDEFGSIND--YSI SPD 60
 38 YRPKSFGIKRMSSENFFRDKDNLYLIFNDNETTEIISNTT1---HNNSSSFTYLED 93

QY	61 GQFILLENNVKQPHSYTASYDIYDLNKRLQITEERIPNNTQWVTWSPVGHKLAYWN 120 94 RKYALLQNYVEKLPHSYTASYHVDIEKEVANEELPNKIQITWSPVGHKLAYWN 153
Db	121 DRYRKIEPNLPSPYRITWTGKEDIYNGITDWWYEEVFSAWSWSPNGTFPLAYAOFND 180 154 NIYKEVPGCSTITTINGEHNKLNGPDWVTEEMPSNTYAWWWSPATSLLAVEFND 23
QY	181 TEPELIEFSYSDESLOQPKTPKTVWVPPAGAVNFTKFEVNTDSISSLNTSQTAP 240 214 TDVVEVTEFSYGENSDQEPHTVTPYKAGARNETVRLFAVNTRSLAVI--NEVEILPP 270
Db	241 ASMLIGDHYLCDVTTWATERISLQWLRLQIYQSYMDICDYDESSGRNCLVAQHIENST 300 271 EELFISIDAYISGNGNWTDKMAVQLRQVNUSLTMCE---GAAMNC-QPYYEQQS 324
QY	301 TGWGRFRPSEPHFTLDGNSFYKTIISNERGYRHYCYFOIDKDDCTFTKGTWVIGIEAL 360 325 TGWGYFQISAPFTQDGKLYKTIISNKGYKTHLFE-GSKDPAVTSGNWETSIAVV 383
QY	361 TSDLYVLSNEYKGMPGGRNLKYKQL-SDPTYKTYCLSCLENPCQYVSVSFSKEAKYQ 419 384 ASNLYVVSNE-GFPGRQLYCLRDGSYSAAQCVENTRCRQOQSAYERSNSKYS 441
Db	420 LRCGPGLPLYTINHSVNDKGRLVLEDNSALDKMLQVQMPSKKLDFILNEKTFWYOMI 479 442 LNCTGPGLPIYTUVVNSNSDNEPTMEEDLKLIDBQMPTEKNSLIIDGELWYQLT 501
QY	480 LPPEFDKSKYPLLDYAGPCSQKADTVFRLNATYLASTENTIVASFDGRGSYQDCK 539 502 LPPEFDKSKYPLLDVGGPGSQKVDOFRLNATYLASTETIVASLDGRSSGYQDCK 561
QY	540 IMHAINRRGTFPEVQDIEARQSKMFGYDUNKRAIAWGWSVCGYTTSMU4GSGCVPKC 599 562 IHCQIYHKLGTLEVQDQTAHKFSSLGFDPKRMATWGWSVCGYTTSMVLSGSGLPKC 621
Db	600 GIAVAPSRWEYTDSSUTERNCLPTPDNLHYRNTSTMRSRNFKOYEVYLHGTA 659 622 GIAVAPSWHYYDSIVTERYMLGPTPDNUYLSSSTYMARAQKFQDVYELHGTA 681
QY	660 NVHFQQSQAIQSKALWDVCFDQAMWYTFEDDHGIASTAHQHTYHMSHFIKQCSL 715 682 NVHFQQAHHISKLVDAQDFETMWYTDKDHG1-GGTANRHITYHMSHFLKQCFNI 736
Db	

RESULT 13

P70092 XENLA
 ID P70092_XENLA PRELIMINARY; PRT; 748 AA.
 AC P70092;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DT Dipeptidyl-peptidase IV (EC 3.4.4.5).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Pipidae; Pipoidea; Xenopus.
 OC Xanopodinae; Xenopus; Xenopus.
 NCBI_TaxID=8355;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Skin.
 RT "cDNA cloning and expression of secreted Xenopus laevis dipeptidyl
 aminopeptidase IV."
 RT Eur. J. Biochem. 247:107-113 (1997).
 DR Vlasak R., Vilas U., Strobl B., Kreil G.,
 DR EMBL; Y08932; CAJ70156.1; -; mRNA.
 DR HSSP; P27487; 1PQ.
 DR MEROPS; S09_003; -.
 DR GO:GO:0016020; C:membrane; IEA.
 DR GO: GO:0004177; F:aminopeptidase activity; IEA.
 DR GO: GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
 DR GO: GO:0016787; F:hydrolyase activity; IEA.

Query Match 59.6%; Score 2309; DB 2; Length 737;
 Best Local Similarity 59.1%; Pred. No. 5e-149; Mismatches 164; Indels 20; Gaps 9;

3 YRLKLYSPRWSIDSEYLYQENNIVLVEAYGNSSVFLENSTPDEFGSIND--YSI SPD 60
 38 YRPKSFGIKRMSSENFFRDKDNLYLIFNDNETTEIISNTT1---HNNSSSFTYLED 93

DR	GO; GO:0006508; P: proteolysis and peptidolysis; IEA.		OS Homo sapiens (Human).
DR	PFam; PF00930; DPPIV N; 1.		OC Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteutharia; Euarchontoglires; Primates; Catarrhini.
InterPro;	IPR002471; Peptidase_S9.		OC Mammalia; Eutheria.
InterPro;	IPR001375; Peptidase_S9.		CC Homo.
InterPro;	IPR002469; Peptidase_S9B.		NCBI_TaxID=9606;
DR	InterPro; IPR000379; Ser_estr.		[1] _;
DR	Pfam; PF00326; Peptidase_S9; 1.		NUCLEOTIDE SEQUENCE.
DR	PROSITE; PS00708; PRO-ENDOPEP_SER; 1.		Cotton M., Maupin R., Hawkins M., Hawkins R.; "The sequence of Homo sapiens BAC clone RP1-5 76116." Submitted (JUN-1999) to the EMBL/CenBank/DDBJ databases.
DR	KW Hydrolase; Transmembrane.		[2] _;
SQ	SEQUENCE 748 AA; 85587 MW; 31CCF61B25104E5B CRC64;		RN [3] _;
Query	Query Match Score 2292.5; DB 2; Length 748;		NUCLEOTIDE SEQUENCE.
Best Local Similarity 58.9%; Fred. No. 6.9e-148;	Mismatches 110; Indels 8		Waterston R. H.; Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
Matches 423; Conservative 110; Mismatches 160; Indels 8	Gaps 11;		RN [4] _;
3 YRLKLYSLRWISDHEVLYKQENNLYVNAEGNSVPLENSTFDEFGHSTIND--YSISPD 60		NUCLEOTIDE SEQUENCE.	
50 YRPKSFLKWKWSENEVFRDINVLIFVNDETTIISNTNTT---HNSNSFTYLS 105		RA Wilson R. K.; Submitted (APR-2005) to the EMBL/GenBank/DDBJ databases.	
61 QFILLENNYTKQRHSYTASYDIDLNRKLITERIPIPNQTWSPVGHCLAYWN 120		RA Wilson R. K.; Submitted (APR-2005) to the EMBL/GenBank/DDBJ databases.	
106 RKY-ALQNYTEKWLWHSYTASYHIDIEKGIVAANELPKIQYITWSPVGHCLAYWN 164		RA Wilson R. K.; Submitted (APR-2005) to the EMBL/GenBank/DDBJ databases.	
121 DIYVKIEPLNLPSYRITWKGDEDIITYGNTIWYEEVEFSAYSALWSPNCFELAYAQAFND 180		DR GO; GO:0016020; -; Genomic_DNA.	
165 NYIREVPGGISTTNGENHKLINGIPDWWYEBMFSTMYALWSPDATSLAVEFND 224		DR GO; GO:0003824; F: catalytic activity; IEA.	
181 TEVPLIEYSFSDESIQLQPKTVRYPKAGAVNPYKEFYVNTDSLSSVTNATSIQTAP 240		DR GO; GO:0004287; F: dipeptidase 1-peptidase IV activity; IEA.	
225 TDVPIYEBSYFSDQSPYHPTVYPKAGRNPYTVLFAVNTSRLAVI --NPyVBLPP 281		DR GO; GO:0004287; F: prollyl oligopeptidase activity; IEA.	
241 ASMLIGDHYLDGVWTQERISLQWIRRIONSYSDIDYDESSGRWNCL--VARQHIEM 298		DR GO; GO:0006508; P: Proteolysis and peptidolysis; IEA.	
282 EBLRSIDHY1SIGINWYTDNQMAVQNLRRIONVSLLTMCE ---GAWNCCOPPVYKQ--- 333		DR InterPro; IPR002471; Pept 59 As.	
299 SITGAWYGRFRPSEPHFTLDGNSFYKLIISNEGYRHICYFOLDKDCTFTIKGTWEVIGIE 358		DR InterPro; IPR001375; Peptidase_S9.	
334 SITGAWYGFQSSAPIFTQDGKIKYKTIISNEGYKTHLFE-GSKDPDVATSGNWVTSIA 392		DR InterPro; IPR002469; Peptidase_S9B.	
359 ALTSDYLYXISNEYKGMPCGRNLKYIQL-SPTYTKTCLSCBLNPERCOYSSFSKEAKY 417		DR Pfam; PF00930; DPPIV_N; 1.	
393 TVASNLFLYYVYNE--GFPGRFLQYK1RDLDSYSQAOCVCTNTRQRCQOYSAFERNISKY 450		DR Pfam; PF00326; Peptidase_S9.	
418 YQLRCRGPGPLDYLTHSSVNDKGRLVLEDNSALDKMLQNYOMPSSKLDFTILNETKFWYQ 477		DR PROSITE; PS00708; PRO-ENDOPEP_SER; 1.	
451 YSLNCNGPGFLPYTYVNNNSNINETRTMEDNEDLKULDEQCMPTKENKSIIDGFELWYQ 510		DR Sequence 760 A; 87713 MW; 7F817B5A4F5142 CRC64;	
478 MILPPHPFDKSKKPLIIDDVAGPCPSQKADTFVRLNWATYLAESTENTIVASFDGRGSQYQG 537		Query Match Score 2145; DB 2; Length 55.3%;	
511 LTLPHPFDKSKKPLIIDDVGPQSKIVDQFVRLNWATYLAESTEKIVASLDGRGSQYQ 570		Best Local Similarity 53.5%; Pred. No. 8.5e-138;	
538 DKIMHAIINRRLGTEFEVDIARARQFSKMGKFVGRNLKQYKTRKQYKTRKQYKTRKQYK 597		Matches 384; Conservative 115; Mismatches 185; Indels 18	
571 DKIMHOIYHKGTLEYQDQITAQHFSSLGFDPKRMATWGSYCGYVTSMLGSGSGLP 630		QY 2 TYRKLYSLRWSIDHEVLYKQ-ENNVLVPAEYGNNSVFLENSTDE	
Db		QY 51 TFSYKTFPFWPSQGELYQDLSQADSVNLTQETGQSYTILSNRTM--	
Qy		Db	59 PDGQTLILENTVQWRSHTASDIYDANKRQLITERIPANTQWV
Db		Qy	107 PDRQPVYLESQSYKWLWRSYSTATYIYDLSNGEFVRGNELPRIQYL
Db		Db	119 NNDIVVKIEPNLPSYRITWKGDEDIYNGITDWWYEEVEYSVASYALW
Qy		Qy	167 QNNIIYLRQGDPPQTITNGRENKINGLPDWYEEVEMATRALKW
Db		Db	179 NDTEVPLIEYSFSDESIQLQPKTVRYPKAGAVNPYKEFYVNTDS
Qy		Qy	227 NDTDIPVIAVSYGCDE-QPYRTNIPVYQPKAGAKNPVVRFLIDITY
Db		Db	239 APASLIGHLYCDVDTWATERISLQWRRIONSYMDICDYEDESG
Db		Qy	282 VPMTIASDXYFSNITTWVDBRVCLQWLRVQNVSYLSICDFREDWQ
Qy		Qy	299 STTGWVGRFRPSEPHFTLDGNSFYKLIISNEGYRHICYFOLDKDCT
Db		Db	342 SRTGAGGGFFVSTPFSIDASHYKIFPSDQGKHHYKDTVENAII
Q537ES_HUMAN PRELIMINARY;	PRT; 760 AA.	QY	359 ALTSDYLYXISNEYKGMPCGRNLKYIQL-SPTYTKTCLSCBLNPERC
ID Q537ES_HUMAN PRELIMINARY;		QY	402 RVTQDSLFISSSEFBEYGRNNTYRISGTYPPSKCKVTCHLKRKC
Q537ES; 13-SEP-2005 (TrEMBL); 31, Created;		Db	418 YQLRSGGPGPLWYHLSYNDKGDRVLEDSALDQMLQNQMPSSKL
DT 13-SEP-2005 (TrEMBL); 31, Last sequence update;		Qy	
DT 13-SEP-2005 (TrEMBL); 31, Last annotation update;		Db	
DE Hypothetical protein FAP.		Qy	
DE			
RESULT 14			
Q537ES_HUMAN			
ID Q537ES_HUMAN PRELIMINARY;			
Q537ES; 13-SEP-2005 (TrEMBL); 31, Created;			
DT 13-SEP-2005 (TrEMBL); 31, Last sequence update;			
DT 13-SEP-2005 (TrEMBL); 31, Last annotation update;			
DE Hypothetical protein FAP.			

Db	462	YALCYGPGPISTLHDGRTDQEIKILEENKELENALKNQLPKEIKKLEDEITLWK	521
Qy	478	MILPPHFEDKSKYKPYLLDYYAGPCSQKADTYVERLNWATYLASTENIIVASFDGRSGCYQ	537
Db	522	MILPQFDRSKYKPYLIQYGGPCSQVRSVFAYNWISLASKCVMIALDGRCTAFQ	581
Qy	538	DKIMHAIRRLGTFEVEDOEAQFSKNGFVDNKRIALGWWSGGYTTSVMSLGGGVF	597
Db	582	DCKLYAVTRKGTVVEVDQTTAVRKPIEFGIDEKRIALGWWSGGYSSALASGTGLF	641
Qy	598	KCGIAVAPVSWEYDSYTERVNLPTPEVDNLHDYRNSTVMSRAENFKOVEYLHGT	657
Db	642	KCGIAVAPVSWEYASVTERFNGLPTKDNLHRYKSTVMAEYFNVDTLLHGT	701
Qy	658	DDNTHFQOSAQISKALVDYGVDFQAMWYTDDEDGTIASSTAHQHYTHMSHFIKQCFSL	715
Db	702	DDNTHFQNSAQIAKALVNAQDVDFQAMWYSDQNHL-SGLSTNHLYTHMTHFLQCFSL	758
RESULT 15			
ID	QBR492_RAT	PRELIMINARY;	PRT; 761 AA.
AC	QBR492;		
DT	01-JUN-2002	(TREMBREL_21, Created)	
DT	01-MAR-2004	(TREMBREL_21, Last sequence update)	
DT	01-MAR-2004	(TREMBREL_26, Last annotation update)	
GN	Name=Fap;		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;		
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathia;		
OC	Muridae; Murinae; Rattus.		
RN	NCBI_TaxID=10116;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RA	ZUCKERSCHWEDT K., Park J. E.;		
RL	Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF493782; AAM11677.1; - mRNA.		
DR	HSSP; P22497; 1PFQ.		
MEROPS;	S09-007; -.		
DR	Ensembl; ENSEMBL0000005679; Rattus norvegicus.		
DR	RGD; 621253; Rap.		
DR	GO; GO:001620; C:membrane; IEA.		
DR	GO; GO:004177; F:aminopeptidase activity; IEA.		
DR	GO; GO:004274; F:dipeptidyl-peptidase IV activity; IEA.		
DR	GO; GO:0016787; F:hydrolase activity; IEA.		
DR	GO; GO:004287; F:prolyl oligopeptidase activity; IEA.		
DR	GO; GO:006508; P:protein binding; IEA.		
InterPro	IPR002471; Pept_S9_AS.		
InterPro	IPR001375; Peptidase_S9.		
InterPro	IPR002465; Peptidase_S9B.		
InterPro	IPR000379; Ser_endopeptidase.		
PFam	PF00930; DPPIV_N; I.		
DR	PF0036; Peptidase_S9; I.		
DR	PROSITE; PS00708; PRO_ENDOPEP_SER; I.		
KW	Hydrolase; Transmembrane.		
SEQUENCE	761 AA; 87843 MW;	EB214BBE4BB14F2C CRC64;	
Query Match Score			
Best Local Similarity 55.2%; Length 761;			
Matches 388; Conservative 54.0%; Pred. No. 1.5e-137;			
Mismatches 189; Indels 13; Gaps 6;			
Qy	2	TYRLKLYSLRWDHHELYKQE-NNIVLNFAEYGNNSVFLENSTPDEFHSIN--DYSIS	58
Db	51	PDGQFILLENYNTYQWRHYSYASDIYDLNKRLQITEERIIPNNTQWVTPSPGHKLAYW	118
Qy	59	PDGQFILLENYNTYQWRHYSYASDIYDLNKRLQITEERIIPNNTQWVTPSPGHKLAYW	118
Db	107	PDGQFILLESYDSKWLWRYSTATYIYD1QNGEVRGVELPR1QYLCSVPSKLAVY	166
Qy	119	NNDDYVKEPLNLPYSRITWTGKEDIIVYNGITDNYEEVPSAYSALWNSPNTFLAYQF	178
Db	167	QNNIYLKORPGDPFQFOITYTGRRNIFGIPDNYEEMLATXALWMSRPGKSYLAVYF	226

THIS PAGE BLANK (USPTO)